

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 6, 2001, 10:53:40 ; Search time 16.05 Seconds
(without alignments)
1775.033 Million cell updates/sec

Title: US-08-878-801-2

Perfect score: 1960
Sequence: 1 MARSLTWRCPCWCLTEDEKA.....VERKDYRDSYLARYLDEINLL 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Maximum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first: 45 summaries

Database: PIR_68:*

1: pirl:1:
2: pirl:2:
3: pirl:3:
4: pirl:4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1960	100.0	374	2 A41096	GTP-binding regula
2	1705	87.0	374	2 B41534	GTP-binding protei
3	1057	53.9	359	2 S43699	GTP-binding regula
4	1046	53.4	353	2 B40891	GTP-binding protei
5	1042	53.2	359	1 RGM50	GTP-binding regula
6	1038	53.0	359	1 RGM511	GTP-binding protei
7	1035	52.8	359	1 S30359	GTP-binding regula
8	1031	52.6	359	1 RGM50	GTP-binding regula
9	1026.5	52.4	359	1 S45700	GTP-binding regula
10	1017.5	51.9	354	2 S33309	GTP-binding regula
11	1014.5	51.8	360	2 JN0115	GTP-binding regula
12	1013	51.7	355	2 A40891	GTP-binding protei
13	1013	51.7	355	2 A41534	GTP-binding protei
14	1013	51.7	355	2 S34347	GTP-binding regula
15	995	50.8	353	2 T15288	GTP-binding regula
16	974	49.7	353	2 T15288	GTP-binding regula
17	787.5	40.2	353	2 S71865	GTP-binding regula
18	766.5	39.1	353	2 T50482	GTP-binding regula
19	762	38.9	354	1 RGFRO2	GTP-binding regula
20	760	38.8	353	1 S25493	GTP-binding regula
21	756.5	38.6	354	1 RGFRO1	GTP-binding regula
22	755	38.5	354	1 S27013	GTP-binding regula
23	752.5	38.4	352	2 B41095	GTP-binding regula
24	750.5	38.3	354	1 T24154	GTP-binding regula
25	748.5	38.2	354	1 S27014	GTP-binding regula
26	748.5	38.2	354	1 S27014	GTP-binding regula
27	746.5	38.1	354	2 S24352	GTP-binding regula
28	746.5	38.1	354	2 S24352	GTP-binding regula
29	746	38.1	377	2 I57490	guanine nucleotide

30	745.5	38.0	354	1 RGH011	GTP-binding regula
31	745.5	38.0	354	1 RGH011	GTP-binding regula
32	745.5	38.0	354	1 RGH011	GTP-binding regula
33	744.5	38.0	354	2 A61035	GTP-binding regula
34	744.5	38.0	354	2 S40508	GTP-binding regula
35	742.5	37.9	354	1 RGH02	GTP-binding regula
36	741.5	37.8	354	1 S28157	GTP-binding regula
37	740.5	37.8	354	2 T19476	GTP-binding regula
38	740.5	37.8	354	1 RGH013	GTP-binding regula
39	740.5	37.8	354	1 RGH02	GTP-binding regula
40	740.5	37.8	354	1 RGH02	GTP-binding regula
41	739.5	37.7	354	1 RGH02	GTP-binding regula
42	739.5	37.7	354	1 RGH02	GTP-binding regula
43	738.5	37.7	354	1 RGH02	GTP-binding regula
44	738.5	37.7	354	1 RGH02	GTP-binding regula
45	737.5	37.6	354	2 S28159	GTP-binding regula

ALIGNMENTS

RESULT 1
A41096
GTP-binding regulatory protein alpha-16 chain - human
C:Species: Homo sapiens (man)
C:Date: 20-Mar-1992 #sequence_rev15ion 20-Mar-1992 #text_change 02-Feb-2001
C:Accession: A41096
R:Amatuda III, T.T.; Steele, D.A.; Slepak, V.Z.; Simon, M.I.
Proc. Natl. Acad. Sci. U.S.A. 88, 5587-5591, 1991
A:Title: Galphal6, a G protein alpha subunit specifically expressed in hematopoietic
A:Reference number: A41096; M01D:91288509
A:Accession: A41096
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-374 <MAS>
A:Cross-references: GB:M63904; NID:g182891; PIDN:AAA35860.1; PID:g182892
C:Superfamily: GTP-binding regulatory protein Gs alpha chain
C:Keywords: GTP binding; nucleotide binding; P-loop
F:49-56/Region: nucleotide-binding motif A (P-loop)
F:277-280/Region: GTP-binding NKXD motif

Query Match	Score	1960	DB 2	Length	374
Best Local Similarity	100.0%				
Matches	374	Conservative	0	Mismatches	0
Indels	0	Gaps	0		
DB	1	MARSLTWRCPCWCLTEDEKAARVDQINIRLLLEOKRQDGEKLLILGPGESGKSTFIK	60		
QY	1	MARSLTWRCPCWCLTEDEKAARVDQINIRLLLEOKRQDGEKLLILGPGESGKSTFIK	60		
DB	1	MARSLTWRCPCWCLTEDEKAARVDQINIRLLLEOKRQDGEKLLILGPGESGKSTFIK	60		
QY	61	QMRITRGAGSEERKGFRLVYONIFVSRAMIEAMERQIPPSRESKHAALVNSOD	120		
DB	61	QMRITRGAGSEERKGFRLVYONIFVSRAMIEAMERQIPPSRESKHAALVNSOD	120		
QY	121	PYKATTEPKRYAAMQWIMDAGIRACYERRRHHLSAVVYSHLERTTEETVPTAQ	180		
DB	121	PYKATTEPKRYAAMQWIMDAGIRACYERRRHHLSAVVYSHLERTTEETVPTAQ	180		
QY	181	DVIRSRMPTTGINEYCFVQKMLRIVDVGOKSERKKWJHCEBNVIALIYLAISYDQ	240		
DB	181	DVIRSRMPTTGINEYCFVQKMLRIVDVGOKSERKKWJHCEBNVIALIYLAISYDQ	240		
QY	241	CLENNQENRMSKSLALFGLILPEWFKSVLILFNKQIILKEKIPTSILATYFPFQ	300		
DB	241	CLENNQENRMSKSLALFGLILPEWFKSVLILFNKQIILKEKIPTSILATYFPFQ	300		
QY	301	PKDAAEAKRFLIDMTYRMTYGVDPGSGSKGARSRLFSHYTCATDTONIRKVFQVR	360		
DB	301	PKDAAEAKRFLIDMTYRMTYGVDPGSGSKGARSRLFSHYTCATDTONIRKVFQVR	360		
QY	361	DSVLARYLDEINLL 374			
DB	361	DSVLARYLDEINLL 374			



Sequence Comparison B3

FEATURES

SOURCE

20 Sep 1964

831 AAACCAACTGGCATCTGTGATGCTGGCGGCCAGAGTCAGAACGTA 880

```

920 CACGAGAGGANTACGACCAGTGCCTGGAGAGIACAAACGAGGAACCCG 969
251 MetlysgIserLeuAlaLeupheGlyTyrTlLeuGluleuPProTyrph 267
970 ATGAGAGAGAGCCCTCGCATGTTGGGACATATCTGGAACTACCCGTGTT 1019
267 elySerThrSerValIleuPheLeuAnlysrThrAspIleLeuGlu 284
1020 CAAGAAGACATCCGTCAATCCTCTTTCACAACAACCGAGATCCTGGAG 1069
284 IuYsIlePProThrSerHisLeuAlaThrTyrTrheProSerPheGIngly 300
1070 AGAAAAACCCACCTCCACCTGGTCAACCAATTCCCCAGTTCCAGGGC 1119
301 ProLySGlnAspAlaGluAlaAlaLysArgPheIleLeuAspMetTyrTh 317
1120 CTAAGAGAGATGCTGAGCGACGCAAGAGATTATCCTGGACATGTACAC 1169
317 rArgMetTyrThrGlyCysValAspGlyProGluGlySerLysGlyYA 334
1170 GAGGATGTACACCGGGGCGCTGGACGGCCCGGCGGCGACAGAAAGGGG 1219
334 IarYsSerIArgArgLeuPheSerHisTyrThrCysAlaThrAspTrGln 350
1220 CAGGATCCCGACGCTTTTCAGCCACTACACNGTGTCCACAGACACAG 1269
351 AspIleArgLysValPheLysAspValArgAsrSerValLeuAlaArgY 367
1270 AACATCGGCAAGGCTTCAAGAGACGCGGGACTCGCTGCTCGCCCGCTA 1319
367 rLeuAspGluIleAsnLeuLeu 374
1320 CCTGGACGAGATCAACCTGCTG 1341

```

seq_name: gb_pr10:HUMGA16

		seq. documentation block:
LOCUS	HIMGAL16	mRNA
DEFINITION	Human G-alpha 16 protein mRNA,	PRI
ACCESSION	M63904	complete cds.
VERSION	M63904.1	GI:183891
KEYWORDS	G-alpha 16 protein.	
SOURCE	Homo sapiens cDNA to mRNA.	
ORGANISM	Homo sapiens	

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
FEATURES
source

1. location/Qualifiers
1. 2060

EnuairPota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Carnivora; Homidae; Homo.
1 (sites)
Amatruda, T.T. III., Steele, D.A., Slepak, V.Z. and Simon, M.I.
G-alpha16, a G protein alpha subunit specifically expressed in
hematopoietic cells
Proc. Natl. Acad. Sci. U.S.A. 88, 5587-5591 (1991)
91288509

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HL-60"
<1..2060
/product="G-alpha-16 protein"
220..1344
/codon_start=1

```

/consensus
 /product="G-alpha16 protein"
 /protein_id="AAA5860.1"
 /db_xref="GI:182892"
 /translation="MARSLIRCCPMCLFEDERAAARVDQENIRLLLEKQKDRGELK
 LLLGPEGSKSTFKKQMLRIIGAGYSEERKRFPLVYQNTIVSRNAIEMERQI
 PFSPESEKHAASLVMSQDPYKVTTFEERAAANQWMLRQVACIERREERHLLDSA
 VYVLSHERTEEGYPTADVLRSRPPTGIGNEYFSVQKTLRIYVDGGQSEKK
 WIRDFEENVALIVLASIENDJCLEENQENRKKSLATGTLITLTPKRSVITFL
 NKMDILEEKIPTSLAIRVPEPSIQGKQDDEAARFIDMTFTYETTCVGDGPGSKGA
 RSRRLSEHYTCARDQNIKRVFQDVDSVYLARIDELINLL"
 631 c 640 g 374 t

ORIGIN

```

alignment_scores:
  Quality: 1960.00      Length: 374
  Ratio: 5.211         Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000

alignment_block:
  US-08-878-801-2 x HMMGA16      ..
  Align seg 1/1 to: HMMGA16 from: 1 to: 2060

```

1	MANAARISerIeuhrThrPArgCysCysProthryPcysIeuThrluAs	17
220	ATGGCCCGCTCGCTGACTGGCCGCTGCTGCCCTTGAGTCCCTGACGAGGA	269
17	poluysaAlaAlaAlaArgValAspGlnIuIleasnArgIleIeuLeug	34
34	IuGlnIuysLysGlnAspArgIuGlnIuIeuysLeuLeuLeuGluPro	50
320	AGCAGAGAGAGACGAGCCCGCGGAGAGCTGAAGCGTGCCTTTGGGCCCA	369
51	GlyGluSerGlyLysSerThrPheIleGlnIuMetArgIleIleHisGlu	67
370	GGCGAGAGCGGGAAGAGACACTTCATCAACAGATGGGATCATCCACG	419
67	YLAAGLYTyrSerGlnGlnIuArgIuysGlyPheArgProIeuValTyrG	84
420	CGCGCGCTACTCGGAGAGAGAGCGCAAGGCTTCCGGCCCTTGCTTACC	469
84	lnAsnIlePheValSerMetArgAlaMetIleGluAlaMetGluArgLeu	100
470	AGAACTCTTCGTGCTCCATGCGGGCCATGATGTCAGGGCCATGAGACGGTG	519
101	GlnIleProPheSerArgProGluSerLysHisHisAlaSerIeuValMe	117
520	CAGATTCATTCAGCGAGCCCGAGACACAGACACACTGACCTGACCTGCAT	569
117	tSerGlnAspProTyrLysValThrThrPheGlnIuLysArgTyrAlaAla	134
570	GAGCCAGAGCCCTTAAAGTGCACAGCTGTGAAGACGTCACGCTGGG	619
134	IaMetGlnIlePheurTPArgAspAlaGlyIleArgAlaCysTyrGluArg	150
620	CCATCAGATGCGCTGTGGAGAGATCGCGGCACTCCGGGCTGCTATGAGCGT	669
151	ArgArgGluPheHisIleuLeuAspSerAlaValTyrTyrIleuSerHisLe	167
670	CGGCGGGAATTCACCTGCTGCTCAATTAGCGCTGTACTACCTGTCCACCT	719
167	uGluArgIleThrGlnGluGlnIuTyrValProThrAlaGlnAspValIeuA	184
720	GGAGGCAATCACCGAGAGAGGCTACTCCCCACAGCTCAGAGACTGCTGC	769
184	rGserArgMetProThrArgIuIleAsnGluTyrCysPheSerValGln	200
770	GCAGCGCATGGCCCACTGGCATCAAGAGATAGTCTTCTCGCTGACG	819
201	LysThrAsnIleuArgIleValAspValGlyGlnIuLysSerGluArgGly	217
820	AAACCAACCTCGGATCTGGAGCGTGGGGGCGAGAAAGTCAGAGCGCTAA	869
217	sLysTrpIleHisCysPheGluAsnValIleAlaIeuIleTyrIleuAlas	234
870	GAATATGATTCATTTTCGAAACGTGATTCGCCCTCATCTACTGCGCT	919
234	erIeuSerGluTyrAspGlnCysLeuGlnIuAsnAsnGlnIuAsnArg	250
920	CACGTAGAGATACGACACAGTCCCTGGAGAGAACAAACCGAGAAACGCG	969

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 6, 2001, 10:53:40 ; Search time 16.05 Seconds
(without alignments)
1775.033 Million cell updates/sec

Title: US-08-878-801-2
 Perfect score: 1960
 Sequence: 1 MARS1TRCCPCWLCUTEDEKA.....VFQDVRSVLARYLDEINLL 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:	219241 seqs, 76174552 residues	219241
matl number of hits satisfying chosen parameters:		219241

Total number of hits satisfying criteria: 1000000
 Minimum DB seq length: 0
 Maximum DB seq length: 20000000000
 Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 hits

```
Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*
```

SUMMARIES

Result No.	Score	Match	Query %	Length	DB	ID	Description
1	1960	100.0	374	2	A41096	GTP-binding regula	
2	1705	87.0	374	2	B41534	GTP-binding protei	
3	1057	53.9	359	2	S45699	GTP-binding regula	
4	1046	53.4	353	2	B40891	GTP-binding protei	
5	1044	53.3	359	1	B40891	GTP-binding regula	
6	1042	53.2	359	2	S71963	GTP-binding protei	
7	1038	53.0	359	1	RGMS11	GTP-binding regula	
8	1035	52.8	359	2	S30359	GTP-binding regula	
9	1031	52.6	359	1	RGHUGY	GTP-binding regula	
10	1026.5	52.4	359	2	S45700	G-alpha-11 protein	
11	1017.5	51.9	354	2	S33309	GTP-binding regula	
12	1014.5	51.8	360	2	JN0115	GTP-binding regula	
13	1013	51.7	355	2	A40891	GTP-binding protei	
14	1013	51.7	355	2	A41534	GTP-binding regula	
15	995	50.8	353	2	S34347	hypothetical prote	
16	974	49.7	353	2	T15288	GTP-binding regula	
17	787.5	40.2	353	2	S71965	G protein alpha ch	
18	766.5	39.1	353	2	T50482	GTP-binding regula	
19	762	38.9	354	1	RGFPO2	GTP-binding regula	
20	760	38.8	353	2	S25493	GTP-binding regula	
21	756.5	38.6	354	1	RGFFO1	GTP-binding regula	
22	755	38.5	354	2	S27013	GTP-binding regula	
23	755	38.4	377	2	B41095	hypothetical prote	
24	752.5	38.4	352	2	T24154	GTP-binding regula	
25	750.5	38.3	384	1	RGXII1	GTP-binding regula	
26	748.5	38.2	354	2	S27014	GTP-binding regula	
27	748.5	38.2	354	2	T50237	GTP-binding regula	
28	746.5	38.1	354	2	S24352	gustducin - rat	
29	746	38.1	377	2	I57490	guanine nucleotide	

ALIGNMENTS

RESULT 1

RESULT 1

A41096

GTP-binding regulatory protein alpha-16 chain - human

C:Species: Homo sapiens (man)

C:Date: 20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change 02-Feb-2001

C:Accession: A41096

C:Amatruda III, T.T.; Steele, D.A.; Slepak, V.Z.; Simon, M.I.
Proc. Natl. Acad. Sci. U.S.A. 88, 5587-5591, 1991

A:title: Galphal6, a G protein alpha subunit specifically expressed in hematopoietic

A:Reference number: A41096; MUID:91388509

A:Accession: A41096

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-374 <RNA>

A:Cross-references: GB:M63804; NID:gl82891; PIDN:AAA35860.1; PID:gl82892

C:Superfamily: GTP-binding regulatory protein Gs alpha chain

C:Keywords: GTP binding; nucleotide binding; P-loop

F:49-56/Region: nucleotide-binding motif A (P-loop)

F:277-280/Region: GTP-binding NKXD motif

Query Match	100.0%;	Score 1960;	DB 2;	Length 374;
Best Local Similarity	100.0%;	Pred. No. 9.2e-147;		
		Mismatches 0;	Indels 0;	Gaps 0;

Matches	3/4:	Conservative
QY	1	MARSLTWRCPCWCLTDEKAAARVDQENIRILLEQQKDQRGELKILLLPGESGKSTFIK 60
Db	1	MARSLTWRCPCWCLTDEKAAARVDQENIRILLEQQKDQRGELKILLLPGESGKSTFIK 60
QY	61	QMRITIHAGYSEEBERKGRPLVYQIFVSMRAMTEAMERLIQIPFSRPESHKHASLVMSQD 120
Db	61	QMRITIHAGYSEEBERKGRPLVYQIFVSMRAMTEAMERLIQIPFSRPESHKHASLVMSQD 120
QY	121	PYKVTTTFEKRYAAAOMWLWRDAGIRACYERRRPFHLSDSAVYYLSHLERTETEGYVPTAQ 180
Db	121	PYKVTTTFEKRYAAAOMWLWRDAGIRACYERRRPFHLSDSAVYYLSHLERTETEGYVPTAQ 180
QY	181	DVLRSRMPPTGTINEYCFVSQVTNLRIVDVGGQKSERKKWHCFENVITALIYLASLSYDQ 240
Db	181	DVLRSRMPPTGTINEYCFVSQVTNLRIVDVGGQKSERKKWHCFENVITALIYLASLSYDQ 240
QY	241	CLEENNOENRMKESIALFGTILELPWFVKSTSVILFLNKTDILEEKIPTSHLATYFPSPFQG 300
Db	241	CLEENNOENRMKESIALFGTILELPWFVKSTSVILFLNKTDILEEKIPTSHLATYFPSPFQG 300
QY	301	PKQDAEAARFFILDMYTRMTYTCVDGPPEGSKKGARSRLFSGHYTCATDTQNIRKFVDVR 360
Db	301	PKQDAEAARFFILDMYTRMTYTCVDGPPEGSKKGARSRLFSGHYTCATDTQNIRKFVDVR 360
QY	361	DSVLARYLDEINLL 374
Db	361	DSVLARYLDEINLL 374

RESULT 2

B41534

GTP-binding protein alpha-15 chain - mouse

C:Species: Mus musculus (house mouse)

C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 02-Feb-2001

C:Accession: B41534

R:Wikie, T.M.; Scherle, P.A.; Strathmann, M.P.; Slepak, V.Z.; Simon, M.I.

Proc. Natl. Acad. Sci. U.S.A. 88, 10049-10053, 1991

A:Title: Characterization of G-protein alpha subunits in the G-q class: expression in mu

A:Reference number: A41534; MUID:92052208

A:Accession: B41534

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-374 <WTL>

A:Cross-references: GB:M80632; NID:g193570; PIDN:AAA37713.1; PID:g193571

C:Superfamily: GTP-binding regulatory protein Gs alpha chain

C:Keywords: GTP binding; nucleotide binding; P-loop

F:49-56/Region: nucleotide-binding motif A (P-loop)

F:277-280/Region: GTP-binding NKXD motif

Query Match 87.0% Score 1705; DB 2; Length 374;

Best Local Similarity 84.8% Pred. No. 11e-126;

Matches 317; Conservative 34; Mismatches 23; Indels 0; Gaps 0;

QY 1 MARSLTWRCPCWCLTEDEKAAARVDQINRLLLEOKKQDRGELKLLLLGPGESGKSTFIK 60

DB 1 MARSLTWRCPCWCLTEDEKAAARVDQINRLLLEOKKQDRGELKLLLLGPGESGKSTFIK 60

QY 61 QMRTIHGAGYSEERKGRFLVYONIPVSMRAMIEMERLQIPFSRPSKHHASLVMSOD 120

DB 61 QMRTIHGAGYSEERKGRFLVYONIPVSMRAMIEMERLQIPFSRPSKHHASLVMSOD 120

QY 121 PKVTFEKEKYAAQWLMDRAGIRACYYRRRHHLLDSAVYLSHLRTEREGYVPTAQ 180

DB 121 PKVTFEKEKYAAQWLMDRAGIRACYYRRRHHLLDSAVYLSHLRTEREGYVPTAQ 180

QY 181 DVLRSRPTTGINEYCFVSKTKRLRVDVGGQSRERKWHICFENVIALIYLSLSEYDQ 240

DB 181 DVLRSRPTTGINEYCFVSKTKRLRVDVGGQSRERKWHICFENVIALIYLSLSEYDQ 240

QY 241 CLEENQENMKESALFTILELPWFKSTSVILFLNKTDILEEKIPTSHLATYFPFSG 300

DB 241 CLEENQENMKESALFTILELPWFKSTSVILFLNKTDILEEKIPTSHLATYFPFSG 300

QY 301 PKODAAKRFILDMYRMTGCVDPGEGSKGARSRLFSHYTCATDTONIRKVKFQVR 360

DB 301 PRDAAKRFILDMYRMTGCVDPGEGSKGARSRLFSHYTCATDTONIRKVKFQVR 360

QY 361 DSVLARYLDEINLL 374

DB 361 DSVLARYLDEINLL 374

RESULT 3

S45699

GTP-binding regulatory protein alpha chain q - African clawed frog

N:Alternative names: G-alpha-q protein

C:Species: Xenopus laevis (African clawed frog)

C:Date: 10-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 02-Feb-2001

C:Accession: S45699

R:Shapira, H.; Way, J.; Lipinsky, D.; Oron, Y.; Battey, J.F.

FEBS Lett. 348, 83-92, 1994

A:Title: Neurokinin B receptor, expressed in Xenopus laevis oocytes, selectively couples

A:Reference number: S45699; MUID:94298961

A:Accession: S45699

A:Molecule type: mRNA

A:Residues: 1-359 <SRA>

A:Cross-references: GB:U10502; NID:g505689; PIDN:AAA52189.1; PID:g505690

C:Superfamily: GTP-binding regulatory protein Gs alpha chain

C:Keywords: GTP binding; nucleotide binding; P-loop; signal transduction

F:46-53/Region: nucleotide-binding motif A (P-loop)

F:274-277/Region: GTP-binding NKXD motif

F:52/Binding site: GTP (Lys) #Status Predicted

F:183/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 53.9% Score 1057; DB 2; Length 359;

Best Local Similarity 56.4% Pred. No. 1e-75;

Matches 204; Conservative 61; Mismatches 85; Indels 12; Gaps 2;

QY 13 CLTDEKAAARVDQINRLLLEOKKQDRGELKLLLLGPGESGKSTFIKQRIIHGAGYSE 72

DB 10 CLSEAKAEARRINDEIERQLRRDKDRARRELLLLLTGTGSGSKSTFIKQRIIHGAGYSD 69

QY 73 EERKGRFLVYONIPVSMRAMIEMERLQIPFSRPSKHHASLVMSODPYKVTTFEKRYA 132

DB 70 EDRKGRFLVYONIPVSMRAMIEMERLQIPFSRPSKHHASLVMSODPYKVTTFEKRYA 129

QY 133 AAMQWLMDRAGIRACYYRRRHHLLDSAVYLSHLRTEREGYVPTAQDVLRSRPTTG 192

DB 130 DAIKYLWMDPGIQECYDRRREYQLSDSTKYVINVDRIATQGYLPTQDVLRSRPTTG 189

QY 193 NEYCFVSKTKRLRVDVGGQSRERKWHICFENVIALIYLSLSEYDQCLLENQENMK 252

DB 190 IEYFEDLQSVIFRVDVGGQSRERKWHICFENVIALIYLSLSEYDQCLLENQENMK 249

QY 253 ESALFGTILELPWFKSTSVILFLNKTDILEEKIPTSHLATYFPFSGPKODAAKRF 312

DB 250 ESKALFTIITYPWFQNSVILFLNKDLLEKIMYSHLVDFYFPYDGPQDAAKAREFI 309

QY 313 LDMYRMTGCVDPGEGSKGARSRLFSHYTCATDTONIRKVKFQDVRDVLARYLDEIN 372

DB 310 LKNEVDL-----NPDSK-----IYSHFTCATDTONIRKVKFQDVRDVLARYLDEIN 357

QY 373 LL 374

DB 358 LV 359

RESULT 4

B40891

GTP-binding protein GL2 alpha chain - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 27-Mar-1992 #sequence_revision 27-Mar-1992 #text_change 02-Feb-2001

C:Accession: B40891

R:Nakamura, F.; Ogata, K.; Shiozaki, K.; Kameyama, K.; Ohara, T.; Nukada, T

J. Biol. Chem. 266, 12676-12681, 1991

A:Title: Identification of two novel GTP-b. nding protein alpha-subunits that lack app

A:Reference number: A40891; MUID:91286303

A:Accession: B40891

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-353 <NAK>

A:Cross-references: GB:D90336; NID:g217567; PIDN:BAAL4350.1; PID:g217568

C:Superfamily: GTP-binding regulatory protein Gs alpha chain

C:Keywords: GTP binding; nucleotide binding; P-loop

F:40-47/Region: nucleotide-binding motif A (P-loop)

F:150-152/Region: GTP-binding SAK/L motif

F:268-271/Region: GTP-binding NKXD motif

Query Match

Best Local Similarity 53.4% Score 1046; DB 2; Length 353;

Matches 203; Conservative 64; Mismatches 83; Indels 12; Gaps 2;

QY 13 CLTDEKAAARVDQINRLLLEOKKQDRGELKLLLLGPGESGKSTFIKQRIIHGAGYSE 72

DB 4 CLSDEVKESKRINAEIERQLRRDKDRARRELLLLLTGTGSGSKSTFIKQRIIHGAGYSE 63

QY 73 EERKGRFLVYONIPVSMRAMIEMERLQIPFSRPSKHHASLVMSODPYKVTTFEKRYA 132

DB 64 EDRKGRFLVYONIPVSMRAMIEMERLQIPFSRPSKHHASLVMSODPYKVTTFEKRYA 123

Db 310 LKMFVDL-----NPDSK-----IIYSHFTCATDENIRFVFAAVKDTILQNLNKEYN 357

Qy 373 LL 374

Db 358 LV 359

RESULT 6

S71963

GTP-binding protein alpha-q - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 02-Feb-2001

C:Accession: S71963

R:Johnson, G. J.; Leis, L.A.; Dunlop, P. C.

Biochem. J. 318, 1023-1031, 1996

A:Title: Specificity of G-alpha(q) and G-alpha(11) gene expression in platelets and e

A:Reference number: S71963; MUID:96433124

A:Accession: S71963

A:Molecule type: mRNA

A:Residues: 1-359 <JOH>

A:Cross-references: EMBL:L76256; NID:g1478071; PID:AA839498.1; PID:g1478072

C:Superfamily: GTP-binding regulatory protein Gs alpha chain

C:Keywords: GTP binding; nucleotide binding; P-loop; signal transduction

F:46-53/Region: nucleotide-binding motif A (P-loop)

F:274-277/Region: GTP-binding NKXD motif

Query Match 53.2%; Score 1042; DB 2; Length 359;

Best Local Similarity 55.8%; Pred. No. 1.5e-74;

Matches 201; Conservative 60; Mismatches 87; Indels 12; Gaps 2;

Qy 13 CLTEDEKAAARVDQENIRLLEOKKODRGELKLLLLPGESGKSTFKOMRIIHGAGYSE 72

Db 10 CLSEEAKEARRINDEIRQLRRDKDARRELKLLLLGTGESGKSTFKQRIIHGSGYSD 69

Qy 73 EERKGRPLVYQNIYFVSMRAMIEMERLQIPFSPRSKHHASLVMSQDPYKVTTFEKRYA 132

Db 70 EDKRGFTKLYQNIYFVSMRAMIEMERLQIPFSPRSKHHASLVMSQDPYKVTTFEKRYA 129

Qy 133 AAMQWLWDAGIRACRYERRERFHLSDSAVYLLSHLEITEGVVPTAODVLRSMPTTGI 192

Db 130 DAISLWDPGICQYDRRREYQLSDSTKYLLNDLRVADPAPLPTQDDVLRVVPVTGI 189

Qy 193 NEYCFVQKTNLRIVDVGQSGKSRKWHCFENVIALIYSLASEYDQCLEENQNRMK 252

Db 190 IEYFDFLQSVIFRVDVGGQSRKWHCFENVIALIYSLASEYDQCLEENQNRMK 249

Qy 253 ESLALFGTILELPWFKSTSVILFNKTDILEEKIPTSHLATYFPFSGPKQDAEAAKRFI 312

Db 250 ESKALFRTIITYPWFQNSVILFNKTDILEEKIPTSHLATYFPFSGPKQDAEAAKRFI 309

Qy 313 LDMYTRMTGCVDPGEGSKKARSRLFSHYTCATDTQNIKRVKFDVRSVLYRLDEIN 372

Db 310 LKMFVDL-----NPDSK-----IIYSHFTCATDENIRFVFAAVKDTILQNLNKEYN 357

RESULT 7

RMS11

GTP-binding regulatory protein G11 alpha chain - mouse

N:Alternate names: guanine nucleotide binding protein G11 alpha chain; heterotrimeric

C:Species: Mus musculus (house mouse)

C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001

C:Accession: B38414

R:Stratmann, M.; Simon, M.I.

Proc. Natl. Acad. Sci. U.S.A. 87, 9113-9117, 1990

A:Title: G protein diversity: a distinct class of alpha subunits is present in verteb

A:Reference number: A38414; MUID:91067657

A:Accession: B38414

A:Molecule type: mRNA

A:Residues: 1-359 <STR>

A:Cross-references: GB:M55411; NID:g193499; PID:AAA63305.1; PID:g193500

C:Comment: The G proteins are a family of guanine nucleotide-binding proteins that re

ains. The beta and gamma chains, required for GTPase activity, appear to be common to

Qy 133 AAMQWLWDAGIRACRYERRERFHLSDSAVYLLSHLEITEGVVPTAODVLRSMPTTGI 192

Db 124 SAITKLWDPGICQYDRRREYQLSDSTKYLLNDLRVADPAPLPTQDDVLRVVPVTGI 183

Qy 193 NEYCFVQKTNLRIVDVGQSGKSRKWHCFENVIALIYSLASEYDQCLEENQNRMK 252

Db 184 IEYFDFLQSVIFRVDVGGQSRKWHCFENVIALIYSLASEYDQCLEENQNRMK 243

Qy 253 ESLALFGTILELPWFKSTSVILFNKTDILEEKIPTSHLATYFPFSGPKQDAEAAKRFI 312

Db 244 ESKALFRTIITYPWFQNSVILFNKTDILEEKIPTSHLATYFPFSGPKQDAEAAKRFI 303

Qy 313 LDMYTRMTGCVDPGEGSKKARSRLFSHYTCATDTQNIKRVKFDVRSVLYRLDEIN 372

Db 304 LKMFVDL-----NPDSK-----IIYSHFTCATDENIRFVFAAVKDTILQNLNKEYN 351

Qy 373 LL 374

Db 352 LV 353

RESULT 5

RMSQ

GTP-binding regulatory protein Gq alpha chain - mouse

N:Alternate names: guanine nucleotide binding protein Gq alpha chain; heterotrimeric G-

C:Species: Mus musculus (house mouse)

C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001

C:Accession: A38414

R:Stratmann, M.; Simon, M.I.

Proc. Natl. Acad. Sci. U.S.A. 87, 9113-9117, 1990

A:Title: G protein diversity: a distinct class of alpha subunits is present in vertebrat

A:Reference number: A38414; MUID:91067657

A:Accession: A38414

A:Molecule type: mRNA

A:Residues: 1-359 <STR>

A:Cross-references: GB:M55412; NID:g193501; PID:AAA63306.1; PID:g193502

C:Comment: The G proteins are a family of guanine nucleotide-binding proteins that relay

ains. The beta and gamma chains, required for GTPase activity, appear to be common to al

C:Superfamily: GTP-binding regulatory protein Gs alpha chain

C:Keywords: GTP binding; heterotrimer; nucleotide binding; P-loop; signal transduction

F:46-53/Region: nucleotide-binding motif A (P-loop)

F:274-277/Region: GTP-binding NKXD motif

F:52/Binding site: GTP (Lys) #status predicted

F:183/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 53.3%; Score 1044; DB 1; Length 359;

Best Local Similarity 55.5%; Pred. No. 1.1e-74;

Matches 201; Conservative 62; Mismatches 87; Indels 12; Gaps 2;

Qy 13 CLTEDEKAAARVDQENIRLLEOKKODRGELKLLLLPGESGKSTFKOMRIIHGAGYSE 72

Db 10 CLSEEAKEARRINDEIRQLRRDKDARRELKLLLLGTGESGKSTFKQRIIHGSGYSD 69

Qy 73 EERKGRPLVYQNIYFVSMRAMIEMERLQIPFSPRSKHHASLVMSQDPYKVTTFEKRYA 132

Db 70 EDKRGFTKLYQNIYFVSMRAMIEMERLQIPFSPRSKHHASLVMSQDPYKVTTFEKRYA 129

Qy 133 AAMQWLWDAGIRACRYERRERFHLSDSAVYLLSHLEITEGVVPTAODVLRSMPTTGI 192

Db 130 DAISLWDPGICQYDRRREYQLSDSTKYLLNDLRVADPAPLPTQDDVLRVVPVTGI 189

Qy 193 NEYCFVQKTNLRIVDVGQSGKSRKWHCFENVIALIYSLASEYDQCLEENQNRMK 252

Db 190 IEYFDFLQSVIFRVDVGGQSRKWHCFENVIALIYSLASEYDQCLEENQNRMK 249

Qy 253 ESLALFGTILELPWFKSTSVILFNKTDILEEKIPTSHLATYFPFSGPKQDAEAAKRFI 312

Db 250 ESKALFRTIITYPWFQNSVILFNKTDILEEKIPTSHLATYFPFSGPKQDAEAAKRFI 309

Qy 313 LDMYTRMTGCVDPGEGSKKARSRLFSHYTCATDTQNIKRVKFDVRSVLYRLDEIN 372

rise; it is specific for each type of G protein.
 C:Superfamily: GTP-binding regulatory protein Gs alpha chain
 C:Keywords: GTP binding; heterotrimer; nucleotide binding; P-loop; signal transduction
 F:46-53/Region: nucleotide-binding motif A (P-loop)
 F:274-277/Region: GTP-binding NKXD motif
 F:52/Binding site: GTP (Lys) #status predicted
 F:183/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 53.0%; Score 1038; DB 1; Length 359;
 Best Local Similarity 55.5%; Pred. No. 3.2e-74;
 Matches 201; Conservative 65; Mismatches 84; Indels 12; Gaps 2;

QY 13 CLTEDEKAARVDQENIRLLLEQKKODRGELKLLLLGPGESGKSTFIKQRIIHGAGYSE 72
 DB 10 CLSDEVKESKRINAEIKQLRRDKRDARRELKLLLLGTGESGKSTFIKQRIIHGAGYSE 69
 QY 73 ERKGFRLVYONIFVSMRAMTEAMERLQIPFSRPSKHHASLVMSQDPYKVTTFEKRYA 132
 DB 70 EDKRGFTKLIVYONIFTAMQSMIRAMETLILYKYEONKANAVLIREVDVEKVTTFEHOYV 129
 QY 133 AAMQWLWDAGIRACRYERREHLLDSAVYVLSHLERITEEGYVPTAODVLRSMPTTGI 192
 DB 130 NAIKTLWSDPGVQECYDRREDFQLSDSAKYVLTVDRIATVGYLPTQODVLRVPTTGI 189
 QY 193 NEYCSVQKTNLRIVDVGQSGKSRKWKHCFFENVALIYLSLEYDQCLEENNOENRMK 252
 DB 190 IEYFPDENLIIIFRMVGVGQSGSRKWKHCFFENVTISIMFLVALSEYDQVLRVSDNENRME 249
 QY 253 ESLALFGTILELPWFKSTSVILFNKTDILEKTIPTSHLATYFPSPQGPQDAEAAKRFI 312
 DB 250 ESKALFRTIITYPWFNSVILFNKDLLEKILYSHLVDFEFGDQPDQAAREFI 309
 QY 313 LDMYTRMTGCVGDPGEGSKGARSRLFSHYTCATDTONIRKVKDVRDVSILARYLDEIN 372
 DB 310 LKMEVDL-----NPDSDK-----IYSHFTCATDTEINIRFVFAAKVDTILQNLKEYN 357
 QY 373 LL 374
 DB 358 LV 359

RESULT 8

GTP-binding regulatory protein Gs alpha chain, phospholipase C-activating - turkey
 N:Alternate names: phospholipase C-activating G protein
 C:Species: Meleagris gallopavo (common turkey)
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 02-Feb-2001
 C:Accession: S30359; S30360
 R:Naurice, D.H.; Waldo, G.L.; Morris, A.J.; Nicholas, R.A.; Harden, T.K.
 Biochem. J. 290, 765-770, 1993
 A:Title: Identification of Galpha(11) as the phospholipase C-activating G-protein of tur
 A:Reference number: S30359; MUID:93207527
 A:Accession: S30359

A:Molecule type: mRNA
 A:Residues: 1-359 <MAW>
 A:Cross-references: GB:X73072; NID:g312254; PIDN:CAA51530.1; PID:g312255
 A:Experimental source: blood
 A:Accession: S30360

A:Molecule type: protein
 A:Residues: 78-92;121-132;158-180;253-256;307-312;339-345;355-359 <MAW>
 A:Experimental source: erythrocytes
 C:Superfamily: GTP-binding regulatory protein Gs alpha chain
 C:Keywords: GTP binding; nucleotide binding; P-loop
 F:46-53/Region: nucleotide-binding motif A (P-loop)
 F:156-158/Region: GTP-binding SAK/L motif
 F:274-277/Region: GTP-binding NKXD motif
 F:52/Binding site: GTP (Lys) #status predicted
 F:183/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 52.8%; Score 1035; DB 2; Length 359;
 Best Local Similarity 55.5%; Pred. No. 5.5e-74;

Matches 201; Conservative 65; Mismatches 84; Indels 12; Gaps 2;
 QY 13 CLTEDEKAARVDQENIRLLLEQKKODRGELKLLLLGPGESGKSTFIKQRIIHGAGYSE 72
 DB 10 CLSDEVKESKRINAEIKQLRRDKRDARRELKLLLLGTGESGKSTFIKQRIIHGAGYSE 69
 QY 73 ERKGFRLVYONIFVSMRAMTEAMERLQIPFSRPSKHHASLVMSQDPYKVTTFEKRYA 132
 DB 70 EDKRGFTKLIVYONIFTAMQSMIRAMETLILYKYEONKANAVLIREVDVEKVTTFEHOYV 129
 QY 133 AAMQWLWDAGIRACRYERREHLLDSAVYVLSHLERITEEGYVPTAODVLRSMPTTGI 192
 DB 130 NAIKTLWSDPGVQECYDRREDFQLSDSAKYVLTVDRIATVGYLPTQODVLRVPTTGI 189
 QY 193 NEYCSVQKTNLRIVDVGQSGKSRKWKHCFFENVALIYLSLEYDQCLEENNOENRMK 252
 DB 190 IEYFPDENLIIIFRMVGVGQSGSRKWKHCFFENVTISIMFLVALSEYDQVLRVSDNENRME 249
 QY 253 ESLALFGTILELPWFKSTSVILFNKTDILEKTIPTSHLATYFPSPQGPQDAEAAKRFI 312
 DB 250 ESKALFRTIITYPWFNSVILFNKDLLEKILYSHLVDFEFGDQPDQAAREFI 309
 QY 313 LDMYTRMTGCVGDPGEGSKGARSRLFSHYTCATDTONIRKVKDVRDVSILARYLDEIN 372
 DB 310 LKMEVDL-----NPDSDK-----IYSHFTCATDTEINIRFVFAAKVDTILQNLKEYN 357
 QY 373 LL 374
 DB 358 LV 359

RESULT 9

GTP-binding regulatory protein Gy alpha chain - human
 N:Alternate names: guanine nucleotide binding protein Gy alpha chain; heterotrimeric
 C:Species: Homo sapiens (man)
 C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
 C:Accession: A39394
 R:Jiang, W.; Pandey, S.; Tran, V.T.; Fong, H.K.W.
 Proc. Natl. Acad. Sci. U.S.A. 88, 3907-3911, 1991
 A:Title: Guanine nucleotide-binding regulatory proteins in retinal pigment epithelial
 A:Reference number: A39394; MUID:91219481
 A:Accession: A39394

A:Molecule type: mRNA
 A:Residues: 1-359 <GIA>
 A:Cross-references: GB:M69013; NID:g183690; PIDN:AAA8624.1; PID:g183691
 C:Comment: The G proteins are a family of guanine nucleotide-binding proteins that re
 ains. The beta and gamma chains, required for GTPase activity, appear to be common to
 rise; it is specific for each type of G protein.
 C:Genetics:

A:Gene: GDB:GN11
 A:Cross-references: GDB:132587; OMIM:139313
 A:Map position: 19p13.3-19p13.3
 C:Superfamily: GTP-binding regulatory protein Gs alpha chain
 C:Keywords: GTP binding; heterotrimer; nucleotide binding; P-loop; signal transductio
 F:46-53/Region: nucleotide-binding motif A (P-loop)
 F:274-277/Region: GTP-binding NKXD motif
 F:52/Binding site: GTP (Lys) #status predicted
 F:183/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 52.6%; Score 1031; DB 1; Length 359;
 Best Local Similarity 55.2%; Pred. No. 1.1e-73;
 Matches 200; Conservative 66; Mismatches 84; Indels 12; Gaps 2;

QY 13 CLTEDEKAARVDQENIRLLLEQKKODRGELKLLLLGPGESGKSTFIKQRIIHGAGYSE 72
 DB 10 CLSDEVKESKRINAEIKQLRRDKRDARRELKLLLLGTGESGKSTFIKQRIIHGAGYSE 69
 QY 73 ERKGFRLVYONIFVSMRAMTEAMERLQIPFSRPSKHHASLVMSQDPYKVTTFEKRYA 132
 DB 70 EDKRGFTKLIVYONIFTAMQSMIRAMETLILYKYEONKANALLIREVDVEKVTTFEHOYV 129

[illegible]

RESULT 10
SS45700
G-alpha-11 protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 10-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 02-Feb-2001
C:Accession: S45700
R:Shapira, H.; Way, J.; Lipinsky, D.; Oron, Y.; Battey, J.F.
FEBS Lett. 348, 89-92, 1994
A:Title: Neuromedin B receptor, expressed in Xenopus laevis oocytes, selectively couples
A:Reference number: S45699; MUID:94298961
A:Accession: S45700
A:Molecule type: mRNA
A:Residues: 1-359 <SHA>
A:Cross-references: GB:U010494; NID:G505687; PID:AAA52188.1; PID:G505688
C:Superfamily: GTP-binding regulatory protein Gs alpha chain
C:Keywords: GTP binding; nucleotide binding; P-loop
F:46-53/Region: nucleotide-binding motif A (P-loop)
F:274-277/Region: GTP-binding NKXD motif

```

Query Match          52.4%; Score 1026.5; DB 2; Length 359;
Best Local Similarity 54.0%; Pred. No. 2.6e-73;
Matches 202; Conservative 65; Mismatches 92; Indels 15; Gaps 3;

1  MARSITWRCPWCLTDEKAAARVDQENIRILLBOKKODRGELKLLLLGGSGSTFIK 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1  MTLSTWAC--CLSEEVKSKRNAETEKOLRRDKKDSRELKLLLLGGSGSTFIK 57
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
61  QMRIITHGAGVSEBERGKRPFLYQNIYFVSMRAMIEMERLIQIPSPESKHAHSVMSOD 120
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
58  QMRIITHGSGVSEEDKGFLLVFQNIFTAQMSIIRAMETLKLYKEQNKANAAQVVREVD 117
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
121 PYKVTTFEKRYAAOMWLWDAGIRACYERRRERFHLLDSAVVYLSHLERITEGYVPTAQ 180
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
118 VBKVCVTECPQYVNAIKNLWSDPGICEYDRRREYQLSDSTKYVITLDVDRISKPGYIPTQ 177
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
181 DVLRSMPITGINCYCFVSOKTNI RIVDVGQKSERKKWHCFENVIALIYSLASLSEYQ 240
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
178 DVLVRVPVPTTGIIIEYFPDENIIFRVMVDVGQSERKKWHCFENVISIMELVASEYQ 237
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
241 CLEENQENRMKESIALFGTILELPWFKSTSVILFLUNKTDILEEKIPTSHLATYFPFSQ 300
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
238 VLVESONEHRMEESKALFTYITYPWQNSVILFLNKKDLLEDKIMYSHLVDFPEFDG 297
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
301 PKQDAEAARKFILDYMTRYMTCGVDPGEGSKKGARSRRLFSHYTCATDTONIRKVFKDVR 360
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
298 PORDAATAREFLIKMFVDL-----NPDSDK-----IYSHFTCATDENTRFVFAAVK 345
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
361 DSVLARYLDEINLL 374
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
346 DTILQHNKLYNLV 359
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 11

S33309 visual Gq alpha chain - northern European squid

S33309 GGRP-binding regulatory protein Gq alpha chain (northern European squid)

C:Species: Loligo forbesi (northern European squid)

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 02-Feb-2001

C:Accession: S33309

R:Ryba, N.J.-P.; Findlay, J.B.C.; Reid, J.D.

C:Accession: S33309

Biochem. J. 292, 333-341, 1993

A:Title: The molecular cloning of the squid (Loligo forbesi) visual Gq-alpha subunit

A:Reference number: S33309; MUID:93277493

A:Accession: S33309

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-354 <RYB>

A:Cross-references: EMBL:L10289

C:Superfamily: GTP-binding regulatory protein Gs alpha chain

C:Keywords: GTP binding; nucleotide binding; P-loop

F:40-47/Region: nucleotide-binding motif A (P-loop)

F:150-152/Region: GTP-binding SAK/L motif

F:269-272/Region: GTP-binding NKXD motif

[illegible]

RESULT: 12
UN0115
GTP-binding regulatory protein dqg alpha chain - fruit fly (*Drosophila melanogaster*)
A:Alternate names: dqg protein
A:Species: *Drosophila melanogaster*
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 02-Feb-2001
C:Accession: JN0115
A:R:Lee, Y.J.; Dobbs, M.B.; Verardi, M.L.; Hyde, D.R.
A:Neuron 5, 889-898, 1990
A:Title: dqg: a *drosophila* gene encoding a visual system-specific G alpha molecule.
A:Reference number: JN0115; MUID:91097801
A:Accession: JN0115
A:Molecule type: DNA
A:Residues: 1-360 <LE>
C:Genetics:
A:Gene: dqg
References: FlyBase:FBgn0004435

[illegible]

RESULT 13
40891
GTP-binding protein G11 alpha chain - bovine
Species: Bos primigenius taurus (cattle)
Date: 27-Mar-1992 #sequence_revision 27-Mar-1992 #text_change 02-Feb-2001
Accession: A40891 Nakamura, F.; Ogata, K.; Shiozaki, K.; Kameyama, K.; Ohara, K.; Haga, T.; Nukada, T.
Biol. Chem. 266:12676-12684, 1991
Title: Identification of two novel GTP-binding protein alpha-subunits that lack apparent
Reference number: A40891
Accession: A40891
Status: preliminary
Molecule type: mRNA
Residues: 1-355 <NAK>
Cross-references: GB:D90335; NID:g217565; PIDN:BAAL4349.1; PID:g217566
Superfamily: GTP-binding regulatory protein CS alpha chain
Keywords: GTP binding; nucleotide binding; P-loop
42-49/Region: nucleotide-binding motif A (P-loop)
132-154/Region: GTP-binding SAK/L motif
170-273/Region: GTP-binding NKXD motif

```

Query Match      51.7%  Score 1013;  DB 2;  Length 355;
Best Local      54.4%  Pred. 2.9e-72;
Matches 199;  Conservative 62;  Mismatches 91;  Indels 14;  Gaps 2;

9  CCPWCLTEDEKAAAVDQEIINRLLLEOKKODRGELKLLLLGPGESGKSPFIKQMRIHGA 68
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
4  CC-CLSAEKEQSRIASIEIQRDRDKDARELKLKLLLGESGKSTFFKQMRIHGS 61
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

QY	69	GYSEERKGRPLVYQNI	EFVSRM	IAEAMER	:QIPFSR	PESKHHASLVNS	ODPKVTT	FE	128
		:		: :	:	: :	:		
Db	62	GSYDEKRGKTLVYQNI	FTAMQMI	RAMDPT	:KIQYVCEQ	KNKAQL	REVEVDK	VSTLS	121
		:		: :	:	: :	:		
QY	129	KRYAAMQWLRDAGI	RACRYER	RRRHFLLDS	AVYYLSH	LERIT	TEEGVYPT	QAQDVLRS	188
		:		: :	:	: :	:		
Db	122	RDQVEAIKQWQPGI	QYDRRRE	YQSDS	KYLT	DIR	AMP	VFPTQODVLR	181
		:		: :	:	: :	:		
QY	189	TTGINEYCFYSQK	NLRIVD	VGQK	SERKKW	HGFEN	VIALIY	LAIS	248
		:		: :	:	: :	:		
Db	182	TTGITTEYFPD	ENIIFRM	VYDVG	QSR	SRKWI	HCF	SVTSLF	241
		:		: :	:	: :	:		
QY	249	NRMKESLAFCTI	LELPWF	KSTVIL	FLNKT	ILEEK	IPIT	SHLATY	308
		:		: :	:	: :	:		
Db	242	NRMESKALFXTI	ITYPWF	LNSVIL	FLNKK	DLEEK	IMY	SHLIS	301
		:		: :	:	: :	:		
QY	309	KRFTILDWYTR	MYTCV	DGPG	SKGARS	RRLF	SHYTCAT	DTNIR	368
		:		: :	:	: :	:		
Db	302	RDFILTKLYQ	-----	-----	-----	-----	-----	-----	368
		:		: :	:	: :	:		
QY	369	DEINLL	374	-----	-----	-----	-----	-----	349
		:		: :	:	: :	:		
QY	350	REFNLV	355	-----	-----	-----	-----	-----	355
		:		: :	:	: :	:		

RESULT
41534

TP-binding protein alpha-14 chain - mouse
 Species: Mus musculus (house mouse)
 Date: 30-Jun-1992 #sequence_revision
 Accession: A41534 #E33833
 Wilkie, T.M.; Scheraga, H.A.; Strathmann, M.P.; Slepak, V.Z.; Simon, M.I.
 Proc. Natl. Acad. Sci. U.S.A. 86, 10049-10053, 1991
 Title: Characterization of G-protein alpha subunits in the G-q class: expression in
 Reference number: A41534; MUID:92052208
 Accession: A41534
 Molecule type: mRNA
 Residues: 1-355 <WIL>
 Cross-references: GB:M80631; NID:g193568; PIDN:AA83322.1; PID:g193569
 Strathmann, M.; Wilkie, T.M.; Simon, M.I.
 Natl. Acad. Sci. U.S.A. 86, 7407-7409 1989
 Title: Diversity of the G-protein family: sequences from five additional alpha-sub
 Reference number: A41534; A33833; MUID:90017488
 Accession: E33833
 Molecule type: mRNA
 Residues: 217-267 <STR>
 Cross-references: GB:M57616; NID:g193380; PIDN:AA63304.1; PID:g193381; GB:M26739
 Keywords: GTP binding; nucleotide binding; protein Gs alpha chain
 42-45/Region: nucleotide-binding motif A (P-loop)
 152-154/Region: GTP-binding SAK/L motif
 270-273/Region: GTP-binding NKXD motif

Query Match	51.7%	Score 1013	DB 2	Length 355
Local Similarity	54.4%	Prod. No. 2,9e-72		
Matches 199	Conservative 61	Mismatches 92	Indels 14	Gaps 2
9	CCPWCCLTDEKAAARVDOENIRLLLEQKQKODRGEJ	KLLLLGPGESGSKSTFFKQWRILHGA	68	
4	CC-CLSAEEKESQISAEIHRVRRDKDARREJ	KLLLLGTGESGSKSTFFKQWRILHGS	61	
69	GYSEBERKGFRLPYQNIYFVSMRAMIEAMERLQIV	FSRPESKHHASHLVMSODPKYKVTTFE	128	
62	GYSDDDRKGFYKLYQNIFTAMQAINRAMDTLRIY	CMCEQNKENAQILREVEVDKVTALS	121	
129	KRYAANMOWLRDAGIRACVYRREFHLLDSAVYI	LSHLERTTPEGVYPTAODYVLRSRMP	198	
122	RDQVAALQWLDPGQECYDRRREYQSDSAKYI	LTDIRIAMPSPFVPTQDDVLRVRVP	181	
189	TTGYNEXCFGSKVNTLRIVDVGGQSKRKWTHCFN	ENVIALIYLASLSEYDQCLLENNOE	248	
182	TTGIIIEYFPDLNIIFRVDVGGQSKRKWTHCFN	SVTSIIFLVALSEYDQVLAECNIE	241	

us-08-878-801-2.rpr

Thu Sep 6 15:16:57 2001

QY 249 NRMKESLALFCTILELPWFKSTSVILEFNKTDILEEKIPTSHLATYFPSPFOGPKQDAEAA 308
 Db 242 NMEESEKALFRTIITYPWFNLSSVILFNKKDLLEEKIMYSHLISYFPEYTGPKQDVKAA 301
 QY 309 KRFILDMYRMVTCVCGPEGSKGARRLFSHYTCATDTONIRKVKFDVRDVSILARYL 368
 Db 302 RDFILKLYQ-----DQNPDKREKVIYSHFTCATDTENIRFVFAAVKDTILQLNL 349
 QY 369 DEINLL 374
 Db 350 REFNLV 355

RESULT 15
 S34347
 GTP-binding regulatory protein Gq alpha chain - great pond snail
 C:Species: *Lymnaea stagnalis* (great pond snail)
 C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 02-Feb-2001
 C;Accession: S65461; S34347
 R;Knol, J.C.; Ramnatsingh, S.; van Kesteren, E.R.; van Minnen, J.; Planta, R.J.; van Hee
 Eur. J. Biochem. 230, 193-199, 1995
 A;Title: Cloning of a molluscan G protein alpha subunit of the Gq class which is expressed
 A;Reference number: S65461; MUID:95324523
 A;Accession: S65461
 A:Molecule type: mRNA
 A:Residues: 1-353 <RNO>
 A:Cross-references: EMBL:233106; NID:g312629; PIDN:CAA80653.1; PID:g312630
 C:Superfamily: GTP-binding regulatory protein Gs alpha chain
 C:Keywords: GTP binding; heterotrimer; nucleotide binding; P-loop; signal transduction
 F:40-48/Region: nucleotide-binding motif A (P-loop)
 F:150-152/Region: GTP-binding NAK/L motif
 F:268-271/Region: GTP-binding NKXD motif
 F:46/Binding site: GTP (Lys) #status predicted
 F:177/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

Query Match 50.8%; Score 995; DB 2; Length 353;
 Best Local Similarity 53.3%; Pred. No. 7.5e-71;
 Matches 193; Conservative 64; Mismatches 93; Indels 12; Gaps 2;

QY 13 CLTEDEKAAARVDQEIINRILLLEOKKQDRGELKLLLLGPGESGKSTFIKQMRIHAGYSE 72
 Db 4 CIPDELKEQKINQEIETEROLKDRKDRRELKLLLLGTGSGSKSTFIKQMRIHAGYSD 63
 QY 73 EERKGRPLVYQNIYFVSMRAMTEAMERLQIPSPESKHASLVMSQDPYKVTTFEKRYA 132
 Db 64 EDRSHIKIYQNIYFAMHAMIRAMDTLNIQINPANRENGNMIRQIDYETVTTFDKPCV 123
 QY 133 AAMQWLWRDAGIRACRYERREHLLDSAVYLYSLHLERITEGYVPTAQDVLRSRMTTGI 192
 Db 124 DAIISLWDDGIQECYDRRREYQLTDSAKYILDSVERISQDYLPFLQDILRVRVPTGI 183
 QY 193 NEYCFVQKTNLRIVDVGQSKERKWIHCFENVIALYSLSEYDQCLENNQENRMK 252
 Db 184 IEYFPDLSIIFRMVDVGQSRERRKWIHCFENVTSIMELVALSEYDQVLVESDNENRME 243
 QY 253 ESLALFCTILELPWFKSTSVILEFNKTDILEEKIPTSHLATYFPSPFOGPKQDAEAAKRFI 312
 Db 244 ESKALFRTIITYPWFQNSVILFNKKDLLEEKIMHSHLVDFPPEFDGPKKEASTAREFI 303
 QY 313 LDNMYRMVTCVCGPEGSKGARRLFSHYTCATDTONIRKVKFDVRDVSILARYLDEIN 372
 Db 304 LKMFVEL-----NPDPK-----IYSHFTCATDTENIRFVFAAVKDTILQLNLKEYN 351
 QY 373 LL 374
 Db 352 LV 353

Search completed: September 6, 2001, 10:54:04
 Job time: 24 sec

Thu Sep 6 15:16:57 2001

us-08-878-801-2.rpr

Page 8

Thu Sep 6 15:16:56 2001

us-08-878-801-2.frag

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 6, 2001, 10:53:40 ; Search time 20.43 seconds
(without alignments)
1109.808 Million cell updates/sec

Title: US-08-878-801-2
Perfect score: 1960
Sequence: 1 MARSLWRCPCWCLTEDEKA.....VFKDVRDVLARYLDELINLL 374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0601: *
1: /SID58/gcgdata/geneseq/geneseq/AA1980.DAT: *
2: /SID58/gcgdata/geneseq/geneseq/AA1981.DAT: *
3: /SID58/gcgdata/geneseq/geneseq/AA1982.DAT: *
4: /SID58/gcgdata/geneseq/geneseq/AA1983.DAT: *
5: /SID58/gcgdata/geneseq/geneseq/AA1984.DAT: *
6: /SID58/gcgdata/geneseq/geneseq/AA1985.DAT: *
7: /SID58/gcgdata/geneseq/geneseq/AA1986.DAT: *
8: /SID58/gcgdata/geneseq/geneseq/AA1987.DAT: *
9: /SID58/gcgdata/geneseq/geneseq/AA1988.DAT: *
10: /SID58/gcgdata/geneseq/geneseq/AA1989.DAT: *
11: /SID58/gcgdata/geneseq/geneseq/AA1990.DAT: *
12: /SID58/gcgdata/geneseq/geneseq/AA1991.DAT: *
13: /SID58/gcgdata/geneseq/geneseq/AA1992.DAT: *
14: /SID58/gcgdata/geneseq/geneseq/AA1993.DAT: *
15: /SID58/gcgdata/geneseq/geneseq/AA1994.DAT: *
16: /SID58/gcgdata/geneseq/geneseq/AA1995.DAT: *
17: /SID58/gcgdata/geneseq/geneseq/AA1996.DAT: *
18: /SID58/gcgdata/geneseq/geneseq/AA1997.DAT: *
19: /SID58/gcgdata/geneseq/geneseq/AA1998.DAT: *
20: /SID58/gcgdata/geneseq/geneseq/AA1999.DAT: *
21: /SID58/gcgdata/geneseq/geneseq/AA2000.DAT: *
22: /SID58/gcgdata/geneseq/geneseq/AA2001.DAT: *

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1960	100.0	374	20	AAV49121
2	1960	100.0	374	21	AAV99841
3	1947	99.3	374	19	AAW42434
4	1711	87.3	374	21	AAV93970
5	1705	87.0	374	20	AAW49120
6	1701	86.8	374	19	AAW42435
7	1037	52.9	1276	20	AAV49129
8	1037	52.9	1394	20	AAV49129
9	1037	52.9	1397	20	AAV49134
10	1037	52.9	1418	20	AAV49131
11	1035.5	52.8	1323	20	AAV49133

12	1033.5	52.7	1303	20	AAV49132	GABA-BR2*Gqo5 fusi
13	1033	52.7	359	20	AAV52705	Human G-alpha-11 p
14	1033	52.7	359	20	AAV29789	Human G-alpha-11 p
15	1029	52.5	359	20	AAV49125	Chimeric Gq15 prot
16	1013	51.7	355	21	AAV15026	Mouse TC-Galpha14
17	746.5	38.1	354	19	AAW74713	Amino acid sequenc
18	746.5	38.1	354	21	AAV59514	Gustatory alpha su
19	746	38.1	377	20	AAV43317	Human G-alpha-13 p
20	745.5	38.0	354	21	AAV85290	Human G-alpha-11 a
21	742.5	37.9	354	14	AAV842424	Rat gustducin alph
22	740	37.8	355	21	AAV85149	Human G-alpha-12 a
23	738.5	37.7	354	21	AAV94207	Human G-alpha-13
24	731.5	37.3	350	14	AAV42426	Bovine rod transdu
25	731.5	37.3	350	21	AAV59536	Bovine rod transdu
26	725.5	37.0	354	15	AAV58924	Alpha subunit of h
27	711.5	36.3	354	14	AAV42425	Bovine cone transd
28	711.5	36.3	354	21	AAV59535	Bovine cone transd
29	707	36.1	381	21	AAV67285	Human G-alpha-12 a
30	692.5	35.3	354	15	AAV58914	Alpha subunit of h
31	654	33.4	394	19	AAW61187	Chimera consisting
32	651	33.2	380	21	AAV23382	Human G-alpha-si p
33	647	33.0	394	19	AAW61186	Chimera consisting
34	646	33.0	394	19	AAW61183	Chimera consisting
35	643	32.8	394	17	AAV94559	Human Gs alpha pro
36	640	32.7	394	19	AAW61188	Wild type G-alpha
37	639	32.6	394	19	AAW61180	Chimera consisting
38	637.5	32.5	394	19	AAW61184	Chimera consisting
39	637.5	32.5	470	20	AAV02227	Chimera consisting
40	637	32.5	394	19	AAW61185	Chimera consisting
41	637	32.5	394	19	AAW61179	Chimera consisting
42	636	32.4	389	19	AAW44948	G-alpha s protein.
43	633	32.3	394	19	AAW61181	Chimera consisting
44	633	32.3	394	19	AAW61181	Chimera consisting
45	622.5	31.8	472	20	AAV02224	Chimeric G-protein

ALIGNMENTS

RESULT 1
AAV49121
ID AAV49121 standard; Protein: 374 AA.

AC AAV49121;

DT 07-JAN-2000 (first entry)

DE G protein alpha 16.

XX G-protein fusion receptor; CaR; calcium receptor; GluR; head injury;
XX metabotropic glutamate receptor; GABABR; chimeric receptor; stroke;
XX gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
XX spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
XX Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
XX cognitive disorder.

OS Homo sapiens.

PN WO9951641-A1.

PD 14-OCT-1999.

XX 02-APR-1999; 99WO-US07333.

XX 03-APR-1998; 98US-0080671.

PA (NPSF) NPS PHARM INC.

XX Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;

PI Simin RT;

XX WPI; 1999-610995/52.

XX

PT New G-protein fusion receptors and chimeras containing domains from
 PT different receptors, used to screen for modulators, potentially useful
 -XX e.g. for treating or preventing stroke or Alzheimer's disease -
 PS Claim 32; Fig 4; 255pp; English.
 XX
 CC The invention relates to G-protein fusion receptors (I) comprising:
 CC (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and
 CC intracellular (ICD) domains, each chosen independently from a CAR
 CC (calcium receptor), GluR (metabotropic glutamate receptor) and GABABR
 CC (gamma-aminobutyric acid receptor); (2) an optional linker attached to
 CC the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the
 CC linker. (1), and recombinant chimeric receptors (CR) without the GP
 CC component, are used to assess function of the various domains and to
 CC identify compounds (e.g. allosteric modulators or antagonists) that act
 CC on these domains. The modulators are potentially useful for treating or
 CC preventing diseases associated with the receptors, e.g. stroke, head or
 CC spinal cord injury, epilepsy, ischemia, hyperparathyroidism, anoxia,
 CC Alzheimer's disease, epilepsies, ischemia, hypoglycemia, anoxia,
 CC disorders and depression. Nucleic acid (II) that encodes (I) is used;
 CC (1) for recombinant production of corresponding proteins; and (2) to
 CC produce cells used in screening for modulators. Use of CAR and mGluR
 CC domains allows presentation of GABABR domains, to a binding agent, in a
 CC form more like the natural domain structure compared with use of
 CC incomplete receptors, lacking one or more domains. By shuffling different
 CC domains, agents can be identified that affect particular domains of a
 CC receptor.
 XX
 SQ Sequence 374 AA;

Query Match 100.0%; Score 1960; DB 20; Length 374;
 Best Local Similarity 100.0%; Pred. No. 1.6e-189;
 Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MARSLTRCCPWCCLTEDEKAAARVDQEIINRILLLEQKKQDRGELKLLLLGPGESGKSTFK 60
 Db 1 marsltwrcpcwcltedekaaarvdqeiinrillleqkkqdrgeklilllpgesgstfk 60
 QY 61 QMRIIHGAGYSEERKGFPLVYQNIIVSVNRAMIEAMERLQIPFSPESKHHASLVMSQD 120
 Db 61 qmrlilhagysseerkgfplvyqnlvsvnrामीearlerlqipfspeskhhaslvmsqd 120
 QY 121 PYKVTTFEKRYAAAMQWLWRDAGIRACYERRRERFHLDSAVYVLSHLERITEGYVPTAQ 180
 Db 121 pykvttfekryaaamqlwrdragiracyerrrrerfhlldsavyvylshleritegyvpta 180
 QY 181 DVLSRSMPTTGINEYCFVQKTNLRIVDVGGQKSERKKWIHCFENVIALIYLASEYDQ 240
 Db 181 dvlsrsmpttgineycfvqktnlrivdvvggkserkwwihcfenvialiyilasseydq 240
 QY 241 CLEENNOENRMKESLALFGTILELPWFKSTSVILFNKTDILEEKIPTSHLATYPPSQG 300
 Db 241 cleennqenrmkeslalfgtilelpwfkstsvilfnktdileekiptshlatypsfg 300
 QY 301 PKQDAEAAKRFILDMYTRMYTCVDPGEGSKKGARSRRLFSHYTCATDTQNIKRVFKDVR 360
 Db 301 pkqdaeaakrfildmytrmytcvdpgegskkgarsrrlfshtcatdtqnirkrvfkdv 360
 QY 361 DSVLARYLDEINLL 374
 Db 361 dsvlaryldeinll 374

RESULT 2

AAAY99841
 ID AAY99841 standard; Protein: 374 AA.
 XX
 AC AAY99841;
 XX
 DT 08-SEP-2000 (first entry)
 XX
 DE Human G-alpha-16 protein.

XX
 KW Human; G-alpha-16; G protein; cytostatic; hyperproliferative disorder;
 KW cancer; inflammation; infection; antisense inhibition.
 XX
 OS Homo sapiens.
 XX
 PN WO200032817-A1.
 XX
 PD 08-JUN-2000.
 XX
 PF 25-AUG-1999; 99WO-US19613.
 XX
 PR 03-DEC-1998; 98US-0205143.
 XX
 PA (ISIS-) ISIS PHARM INC.
 XX
 PI Cowser LM;
 DR WPI: 2000-412354/35.
 DR N-PSDB; AAM48751.
 XX
 PT A new antisense compound for inhibiting the expression of human
 PT G-alpha-16 and treating, preventing or delaying infections,
 PT inflammation or hyperproliferative disorders such as cancer -
 XX
 PS Example 10; Page 82-84; 100pp; English.
 XX

CC The present sequence is the human G-protein G-alpha-16. G-alpha-16
 CC interacts differentially with several receptor types including members of
 CC the opioid and chemokine receptor families. Antisense oligonucleotides
 CC that target the nucleotide sequence encoding the present protein can be
 CC used to modulate the expression of G-alpha-16. They may be used to
 CC inhibit the expression of G-alpha-16 in human cells and tissues and thus
 CC to treat disorders associated with G-alpha-16, such as hyperproliferative
 CC disorders, especially cancer. Infections, inflammation or tumour
 CC formation can be prevented or delayed. The compounds can be
 CC used in research and diagnostics in sandwich and other assays.
 XX
 SQ Sequence 374 AA;

Query Match 100.0%; Score 1960; DB 21; Length 374;
 Best Local Similarity 100.0%; Pred. No. 1.6e-189;
 Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MARSLTRCCPWCCLTEDEKAAARVDQEIINRILLLEQKKQDRGELKLLLLGPGESGKSTFK 60
 Db 1 marsltwrcpcwcltedekaaarvdqeiinrillleqkkqdrgeklilllpgesgstfk 60
 QY 61 QMRIIHGAGYSEERKGFPLVYQNIIVSVNRAMIEAMERLQIPFSPESKHHASLVMSQD 120
 Db 61 qmrlilhagysseerkgfplvyqnlvsvnrामीearlerlqipfspeskhhaslvmsqd 120
 QY 121 PYKVTTFEKRYAAAMQWLWRDAGIRACYERRRERFHLDSAVYVLSHLERITEGYVPTAQ 180
 Db 121 pykvttfekryaaamqlwrdragiracyerrrrerfhlldsavyvylshleritegyvpta 180
 QY 181 DVLSRSMPTTGINEYCFVQKTNLRIVDVGGQKSERKKWIHCFENVIALIYLASEYDQ 240
 Db 181 dvlsrsmpttgineycfvqktnlrivdvvggkserkwwihcfenvialiyilasseydq 240
 QY 241 CLEENNOENRMKESLALFGTILELPWFKSTSVILFNKTDILEEKIPTSHLATYPPSQG 300
 Db 241 cleennqenrmkeslalfgtilelpwfkstsvilfnktdileekiptshlatypsfg 300
 QY 301 PKQDAEAAKRFILDMYTRMYTCVDPGEGSKKGARSRRLFSHYTCATDTQNIKRVFKDVR 360
 Db 301 pkqdaeaakrfildmytrmytcvdpgegskkgarsrrlfshtcatdtqnirkrvfkdv 360
 QY 361 DSVLARYLDEINLL 374
 Db 361 dsvlaryldeinll 374

RESULT 3
ID AAW42434 standard; Protein; 374 AA.
XX AAW42434;
AC AAW42434;
DT 22-JUN-1998 (first entry)
XX Murine promiscuous G-alpha 16 protein.
XX Promiscuous G-alpha 16 protein; mouse; G-protein coupled receptor;
KW ligand; agonist; antagonist; signal transduction.
XX Mus musculus.
XX WO9748820-A1.
XX 24-DEC-1997.
XX 20-JUN-1997; 97WO-US10642.
XX 21-JUN-1996; 96US-0020234.
XX (AURO-) AURORA BIOSCIENCES CORP.
PA Negulescu PA, Offermanns S, Simon M, Zuker C;
PI WPI; 1998-063158/06.
DR N-PSDB; AAV03464.
XX Stable cells containing sequence encoding promiscuous G-alpha
PT protein - useful to identify G-protein coupled receptors or ligands,
PT and agonists or antagonists of signal transduction in cells
XX Disclosure; Page 49-53; 72pp; English.
XX This protein comprises murine G-alpha 16 protein, a promiscuous
CC G-protein whose subunits allow coupling with G-protein coupled
CC receptors (GPCRs) that normally couple with G-proteins of other
CC families. Stable isolated cells are claimed that include a
CC construct comprising an inducible promoter linked to a nucleic
CC acid encoding a promiscuous G-alpha protein. A murine G-alpha 16
CC protein polynucleotide (see AAV03464) or a human G-alpha 15 protein
CC polynucleotide (see also AAV03465) is preferably used. The cells
CC may also comprise a second construct in which a reporter gene is
CC linked to a second promoter that is modulated by a promiscuous
CC G-alpha protein. The cells are used in novel methods for
CC identifying a GPCR for a given ligand, or vice versa, or modulators
CC of signal transduction in a cell and for classifying ligands as
CC agonists or antagonists. Since live cells are used, any identified
CC receptor or ligand can be cloned, and use of fluorescent detection
CC permits characterization of individual cells.
XX Sequence 374 AA;
SQ
Query Match 99.3%; Score 1947; DB 19; Length 374;
Best Local Similarity 99.2%; Pred. No. 3.3e-186;
Matches 371; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 MARSLTWCCPCLTDEKAAARVDQENRILLEOKKQDRGELKLLLPGESGKSTFK 60
DB 1 marsltwccpcltdekaarvdqenrilleqkkgdrgelklllpgesgkstfk 60
QY 61 QMRITHGAGYSEERKGRFLPYQNYFVSMRAMIFAMERLQIPFSRPSKHHASLVMSQD 120
DB 61 qmriihgagyseerkgrflpyqnyfvsmramieamerlqipfsrpskhhaslvmqsd 120
QY 121 PKYVTFEKKRYAAAMQWLWRDAGIACVRRERREFHLLDSAVYLLSHLERITEGVPTAQ 180
DB 121 pykvttfekkryaaamqwlwrdagiacvrrerrefhlldsavylshleritegvptaq 180

QY 181 DVLRSRMPPTGINEXCFVSQKTNLRIVDVGQKSERKKWIHCFENVIALIYLASLSEYDQ 240
DB 181 dvlrsrmpptginexcfvsqktnlrivdvggrserkkwi hcfenvialiy laslseydq 240
QY 241 CLEENNOENRMKESALFCTILELFWFKSTSVILFLNKTDILLEEKIPTSHLATYFFSFQ 300
DB 241 cleennoenrmkesal fctilelfpwfkstsvilflnkt dilleekiptshlatyffsfq 300
QY 301 PKODAEAAKRFILDMYRMVTCVDPGEGSKGARSRRFLFSHYTCATDTQNIKRVFKDVR 360
DB 301 pkodaeeakrfildmyrmvtcvcvegs kkgarsrrlfs hytcatdtqnikrvfkdv r 360
QY 361 DSVLARYLDEINLL 374
DB 361 dsvlaryldeinll 374
RESULT 4
AA93970
ID AAY93970 standard; Protein; 374 AA.
XX AAY93970;
AC AAY93970;
XX 03-OCT-2000 (first entry)
DT Amino acid sequence of a rabbit G-protein alpha 16 polypeptide.
DE Rabbit; G-protein alpha 16; signal transduction; Immunogen.
KW Oryctolagus cuniculus.
XX WO200036920-A1.
PN 29-JUN-2000.
PD 21-DEC-1999; 99WO-US30726.
XX 22-DEC-1998; 98US-0218489.
XX (SMIK) SMITHKLINE BEECHAM CORP.
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX Ames RS, Feild JA, Testa T;
PI WPI; 2000-442490/38.
DR N-PSDB; AAA57483.
XX Novel rabbit G-alpha 16 polypeptides for antibody production and for
PT screening antagonists and agonists of G-coupled protein receptor -
XX Claim 1; Page 26-27; 30pp; English.
XX The present sequence represents a rabbit G-protein alpha 16 polypeptide.
CC The polypeptide is believed to be a member of the GTP-binding regulatory
CC protein family. The polypeptide is a regulator of signal transduction.
CC The polypeptide is useful for treating an individual in need of
CC enhanced activity or expression of G-protein alpha 16. Antagonists are
CC useful for treating an individual in need to inhibit the activity or
CC expression of G-protein alpha 16. The polypeptide is also useful as an
CC immunogen to produce antibodies and for producing immunological response
CC in mammals. The G-protein alpha 16 polypeptides and polynucleotides are
CC useful in screening for antagonists and agonists of G-coupled protein
CC receptors.
XX Sequence 374 AA;
SQ
Query Match 87.3%; Score 1711; DB 21; Length 374;
Best Local Similarity 85.8%; Pred. No. 2.4e-164;
Matches 321; Conservative 29; Mismatches 24; Indels 0; Gaps 0;
QY 1 MARSLTWCCPCLTDEKAAARVDQENRILLEOKKQDRGELKLLLPGESGKSTFK 60
DB 1 marsltwccpcltdekaarvdqenrilleqkkgdrgelklllpgesgkstfk 60

Db 1 marslawrcpwcldesekaaarvdqgeitrllehhrrqvigelklilllgtgesgkstfik 60
 -Qy 61 QMRIHAGYSEERKGRPLVQNTFVSMRAMTEAMERLQIPFSPESKHASLVMSOD 120
 Db 61 qmrihagayseerdkgrplvqntfvlsmraamdrilqipsrpeskhaslvmsq 120
 Qy 121 PYKVTFEYAAAMQWLWDAGIRACYERRRHHLLDSAVYLSHLERITEEGYVPTAQ 180
 Db 121 pykvtfeyalavaglsrwdagiracyerrrhhllsdavylshleriteegypvac 180
 Qy 181 DVLRSMPPTGNEYCFVQKTNLRIVDVGQSKSERKWHCPNVTALYLASLSEYDQ 240
 Db 181 dvlrsmpptgineycfvqktnlrivdvgqskserkwhcfnvtdlcalifiaslsydg 240
 Qy 241 CLEENQENRMKESLALFGTILPFWPKSTSVILFLNKTDLIEEKIPTSHLATYFPSPQ 300
 Db 241 cleengqenrmqeslalfgtvlpfwpratsvflfntdiledkvtvrtshlatyfpgr 300
 Qy 301 PKQDAEAAKRFILDMYTRMYTCVGDGPEGSGKSGARSRLFSHYTCATDTQNIKRVKDV 360
 Db 301 pkqdaeaakrfilelytryvagaagdgsgkgsrslfshytcatdtgnirkvfkdv 360
 Qy 361 DSVLARYLDEINLL 374
 Db 361 dsvlaryldeinll 374

RESULT 5
 AAV49120
 ID AAV49120 standard; Protein; 374 AA.
 AC AAV49120;
 XX
 DT 07-JAN-2000 (first entry)
 DE G protein alpha 15.
 XX G-protein fusion receptor; Ca; calcium receptor; GluR; head injury;
 KW metabotropic glutamate receptor; GABAR; chimeric receptor; stroke;
 KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
 KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
 KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
 KW cognitive disorder.
 OS Homo sapiens.
 XX
 PN WO9951641-A1.
 XX
 PD 14-OCT-1999.
 XX
 PF 02-APR-1999; 99WO-US07333.
 XX
 PR 03-APR-1998; 98US-0080671.
 XX
 PA (NPSF-) NPS PHARM INC.
 XX
 PI Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
 PI Simin RT;
 XX
 DR WPI; 1999-610995/52.
 XX
 PT New G-protein fusion receptors and chimeras containing domains from
 PT different receptors, used to screen for modulators, potentially useful
 PT e.g. for treating or preventing stroke or Alzheimer's disease -
 XX
 PS Claim 32; Fig 4; 255pp; English.
 XX
 CC The invention relates to G-protein fusion receptors (I) comprising:
 CC (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and
 CC intracellular (ICD) domains, each chosen independently from a Car
 CC (calcium receptor), GluR (metabotropic glutamate receptor) and GABAR
 CC (gamma-aminobutyric acid receptor); (2) an optional linker attached to
 CC the C-terminus of ICD; and (3) a G-protein (Gp) linked to ICD or the

CC linker. (I), and recombinant chimeric receptors (CR) without the Gp
 CC component, are used to assess function of the various domains and to
 CC identify compounds (e.g. allosteric modulators or antagonists) that act
 CC on these domains. The modulators are potentially useful for treating or
 CC preventing diseases associated with the receptors, e.g. stroke, head or
 CC spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anoxia,
 CC Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive
 CC disorders and depression. Nucleic acid (II) that encodes (I) is used;
 CC (1) for recombinant production of corresponding proteins; and (2) to
 CC produce cells used in screening for modulators. Use of Car and mGluR
 CC domains allows presentation of GABAR domains, to a binding agent, in a
 CC form more like the natural domain structure compared with use of
 CC incomplete receptors, lacking one or more domains. By shuffling different
 CC domains, agents can be identified that affect particular domains of a
 CC receptor.
 CC
 XX Sequence 374 AA;
 SQ

Query Match 87.0%; Score 1705; DB 20; Length 374;
 Best Local Similarity 84.8%; Pred. No. 9.5e-164;
 Matches 317; Conservative 34; Mismatches 23; Indels 0; Gaps 0;

Qy 1 MARSILTWRCCPWCCLTDEKAAARVDQGEITRLLEHRRQVIGELKILLIGTESGKSTFIK 60
 Db 1 marsiltwrcpwcclteeektaaridqeintrilleqkqereelkllilgpgesgkstfik 60

Qy 61 QMRIHAGYSEERKGRPLVQNTFVSMRAMTEAMERLQIPFSPESKHASLVMSOD 120
 Db 61 qmrihagayseerdkgrplvqntfvlsmraamdrilqipsrpeskhaslvmtqd 120

Qy 121 PYKVTFEYAAAMQWLWDAGIRACYERRRHHLLDSAVYLSHLERITEEGYVPTAQ 180
 Db 121 pykvtfekypavagylwrdagiracyerrrhhllsdavylshleriteegysiptaq 180

Qy 181 DVLRSMPPTGNEYCFVQKTNLRIVDVGQSKSERKWHCPNVTALYLASLSEYDQ 240
 Db 181 dvlrsmpptgineycfvqktnlrivdvgqskserkwhcfnvtdlcalifiaslsydg 240

Qy 241 CLEENQENRMKESLALFGTILPFWPKSTSVILFLNKTDLIEEKIPTSHLATYFPSPQ 300
 Db 241 cleengqenrmeslalfgtvlpfwktsvflfntdiledkhtshlatyfpfsg 300

Qy 301 PKQDAEAAKRFILDMYTRMYTCVGDGPEGSGKSGARSRLFSHYTCATDTQNIKRVKDV 360
 Db 301 prdaeaakrfildmyarvyascaepqdggrkgsrarrffahrfcatdtqsvrsvfkdv 360

Qy 361 DSVLARYLDEINLL 374
 Db 361 dsvlaryldeinll 374

RESULT 6
 AAV42435
 ID AAV42435 standard; Protein; 374 AA.
 XX
 AC AAV42435;
 XX
 DT 22-JUN-1998 (first entry)
 XX
 DE Human promiscuous G-alpha 15 protein.
 XX
 KW Promiscuous G-alpha 15 protein; human; G-protein coupled receptor;
 KW ligand; agonist; antagonist; signal transduction.
 XX
 OS Homo sapiens.
 XX
 PN WO9748820-A1.
 XX
 PD 24-DEC-1997.
 XX
 PF 20-JUN-1997; 97WO-US10642.
 XX

PR 21-JUN-1996; 96US-0020234.

XX (AURO-) AURORA BIOSCIENCES CORP.

XX Negulescu PA, Offermanns S, Simon M, Zuker C;

XX WPI; 1998-063158/06.

DR N-PSDB; AAV03465.

XX Stable cells containing sequence encoding promiscuous G-alpha protein - useful to identify G-protein coupled receptors or ligands, and agonists or antagonists of signal transduction in cells

PS Disclosure; Page 46-49; 72pp; English.

XX This protein comprises human G-alpha 15 protein, a promiscuous G-protein whose subunits allow coupling with G-protein coupled receptors (GPCRs) that normally couple with G-proteins of other families. Stable isolated cells are claimed that include a construct comprising an inducible promoter linked to a nucleic acid encoding a promiscuous G-alpha protein. A human G-alpha 15 protein polynucleotide (see AAV03465) or a mouse G-alpha 16 protein polynucleotide (see also AAV03464) is preferably used. The cells may also comprise a second construct in which a reporter gene is linked to a second promoter that is modulated by a promiscuous G-alpha protein. The cells are used in novel methods for identifying a GPCR for a given ligand, or vice versa, or modulators of signal transduction in a cell and for classifying ligands as agonists or antagonists. Since live cells are used, any identified receptor or ligand can be cloned, and use of fluorescent detection permits characterisation of individual cells.

XX Sequence 374 AA;

Query Match 86.8%; Score 1701; DB 19; Length 374;

Best Local Similarity 84.5%; Pred. No. 2.4e-163;

Matches 316; Conservative 35; Mismatches 23; Indels 0; Gaps 0;

QY 1 MARSLTRCCPWCILTEDEKAAARVDOETNRLLEQKQDRCGLKILLGPGESCKSTFIK 60

Db 1 marsltwgccpwcilteektaaridqelnrlleqkqgeelkillingesgskstfik 60

QY 61 QMRIIHGAGYSEERKGRPLVYONIFVSMRAMIEMERLQIPFSRPSKHSHSLVMSQD 120

Db 61 qmriihgvyseedrfrlliynqifvsmqamidamdrlqipfsrpskqhaslviqtd 120

QY 121 PYKVTTEKRYAAMQWLWRDAGTRACYERRRPHLLDSAVYVLSHLRITTEGVPTAQ 180

Db 121 pykvstfekpyavamqylwrddagtracyerrrefhlldsavyvylshlerisedsviptaq 180

QY 181 DVLRSRMPPTTGINCYCFVQKTNLRIVDVGQSKSRKWKHCFENVIALIYLSLSEYDQ 240

Db 181 dvlrsrmpptgincycfvsktklrvdvgggrseerkrwhcfenvialiyaslseydq 240

QY 241 CLEENQENRMKESLALFTGITLELPWFKSTSVILFLNKTDLERKIPTSHTATYFPSPQ 300

Db 241 cleendqenmeeslalfstillelpwfkstsvilflnktldiledkihtshlatyfpsqg 300

QY 301 PKQDAEAAKRIIDMYRMYTGCVDGEGSKGARSRLFSHYTCATDQNIKRVFKDVR 360

- Db 301 prrdaeaakfildmyarvyascaepdgggrkgrarrfahftcoatdqtgsrvsfkdv 360

QY 361 DSVLARYLDEINLL 374

Db 361 dsvlaryldeinll 374

RESULT 7

AAV49127

ID AAV49127 standard; Protein; 1276 AA.

XX

AC AAV49127;

XX

DT

XX

DE

XX

XX

KW

KW

KW

KW

KW

XX

OS

XX

XX

PN

XX

XX

PD

XX

PF

XX

XX

PR

XX

XX

PA

XX

PI

XX

PI

XX

DR

XX

DR

XX

XX

PT

XX

PT

XX

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

07-JAN-2000 (first entry)

phCar/hmGluR2*Gq15 fusion construct protein sequence.

G-protein fusion receptor; CaR; calcium receptor; GluR; head injury; metabotropic glutamate receptor; GABABR; chimeric receptor; stroke; gamma-aminobutyric acid receptor; allosteric modulator; antagonist; spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia; Alzheimer's disease; hyperparathyroidism; osteoporosis; depression; cognitive disorder.

XX Homo sapiens.

XX WO9951641-A1.

XX 14-OCT-1999.

XX 02-APR-1999; 99WO-US07333.

XX 03-APR-1998; 98US-0080671.

XX (NPSP-) NPS PHARM INC.

XX Stormann TM, Hammerland LG, Storzjohann LL, Busby JG, Garrett JE;

XX Simin RT;

XX WPI; 1999-610995/52.

XX N-PSDB; AAZ31058.

XX New G-protein fusion receptors and chimera containing domains from

XX different receptors, used to screen for modulators, potentially useful

XX e.g. for treating or preventing stroke or Alzheimer's disease -

XX Example 1; Fig 12; 255pp; English.

XX The invention relates to G-protein fusion receptors (1) comprising:

XX (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and

XX intracellular (ICD) domains, each chosen independently from a CaR

XX (calcium receptor), GluR (metabotropic glutamate receptor) and GABABR

XX (gamma-aminobutyric acid receptor); (2) an optional linker attached to

XX the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the

XX linker. (1), and recombinant chimeric receptors (CR) without the GP

XX component, are used to assess function of the various domains and to

XX identify compounds (e.g. allosteric modulators or antagonists) that act

XX on these domains. The modulators are potentially useful for treating or

XX preventing diseases associated with the receptors, e.g. stroke, head or

XX spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anoxia,

XX Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive

XX disorders and depression. Nucleic acid (11) that encodes (1) is used:

XX (1) for recombinant production of corresponding proteins; and (2) to

XX produce cells used in screening for modulators. Use of CaR and mGluR

XX domains allows presentation of GABABR domains, to a binding agent, in a

XX form more like the natural domain structure compared with use of

XX incomplete receptors, lacking one or more domains. By shuffling different

XX domains, agents can be identified that affect particular domains of a

XX receptor.

XX Sequence 1276 AA;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Query Match 52.9%; Score 1037; DB 20; Length 1276;

Best Local Similarity 55.4%; Pred. No. 1.9e-95;

Matches 200; Conservative 61; Mismatches 88; Indels 12; Gaps 2;

QY 13 CLTEDEKAAARVDOENRLLLEQKQDRCGLKILLGPGESCKSTFIKQRIIHGAGYSE 72

Db 927 clseakearrindelegrldrdrarrelklilgtgesgskstfikqmrllhsgysd 986

QY 73 PERKGRPLVYONIFVSMRAMIEMERLQIPFSRPSKHSHSLVMSQDQPKVTTTEKRYA 132

Db 987 edkrgftklvyniftamqamiramndtikipykqhnhkaqlvrevdvksafenpyv 1046

QY 133 AAMQWLWDGIRACRYERREFFHLLDSAVYVLSHLERITEGYVPTAQDVLSRMPPTGI 192
 Db 1047 daikslwvdlqecydrreyqlsdsktyllndlradvapylptqdvrvrvptgti 1106
 QY 193 NEYCFVSQVNTLRIVDVGQSKSRKWHCFENVALIYASLSEYDQCLEENQENRMK 252
 Db 1107 leypfdlqsvifrmvvggrsrrkwhcfenvtslmflvalseydvqlvivesdnenme 1166
 QY 253 ESALFPGTILELPWFKSTSVILEFNKTDILEEKIPTSHLATYPPSFGPKQDAEAAKRFI 312
 Db 1167 eskalftrltiypwfnqssvllfnkdlleeklmshlyvdyfpeydgprdaqarefi 1226
 QY 313 LDMYTRMTGCVDPGSGKSGKARSRLFSHYTCATDTQNIKRVKDVRSVLAARYLDEIN 372
 Db 1227 lkmfvd-----npdsdk-----liyshtcatdenirfvfaavkdtlqnlkdcg 1274
 QY 373 L 373
 Db 1275 L 1275
 RESULT 8
 AAY49129
 ID AAY49129 standard; Protein: 1394 AA.
 AC AAY49129;
 XX
 XX 07-JAN-2000 (first entry)
 DE pmGluR2/Car*Galpha15 fusion construct protein sequence.
 KW G-protein fusion receptor; Car; calcium receptor; GluR; head injury;
 KW metabotropic glutamate receptor; GABABR; chimeric receptor; stroke;
 KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
 KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
 KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
 KW cognitive disorder.
 XX
 OS Homo sapiens.
 XX WO9951641-A1.
 PN 14-OCT-1999.
 XX
 XX 02-APR-1999; 99WO-US07333.
 PF 03-APR-1998; 98US-0080671.
 PR (NPSF-) NPS PHARM INC.
 PA Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
 PI Simin RT;
 XX
 DR WPI; 1999-610995/52.
 DR N-PSDB; AA231060.
 XX
 XX New G-protein fusion receptors and chimeras containing domains from
 PT different receptors, used to screen for modulators, potentially useful
 PT e.g. for treating or preventing stroke or Alzheimer's disease -
 XX
 PS Example 1; Fig 12; 255pp; English.
 XX
 XX The invention relates to G-protein fusion receptors (I) comprising:
 CC (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and
 CC intracellular (ICD) domains, each chosen independently from a Car
 CC (calcium receptor), GluR (metabotropic glutamate receptor) and GABABR
 CC (gamma-aminobutyric acid receptor); (2) an optional linker attached to
 CC the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the
 CC linker. (1), and recombinant chimeric receptors (CR) without the GP
 CC component, are used to assess function of the various domains and to
 CC identify compounds (e.g. allosteric modulators or antagonists) that act
 CC on these domains. The modulators are potentially useful for treating or
 CC preventing diseases associated with the receptors, e.g. stroke, head or

CC spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anoxia,
 CC Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive
 CC disorders and depression. Nucleic acid (II) that encodes (1) is used:
 CC (1) for recombinant production of corresponding proteins; and (2) to
 CC produce cells used in screening for modulators. Use of Car and mGluR
 CC domains allows presentation of GABABR domains, to a binding agent, in a
 CC form more like the natural domain structure compared with use of
 CC incomplete receptors, lacking one or more domains. By shuffling different
 CC domains, agents can be identified that affect particular domains of a
 CC receptor.
 XX
 SQ Sequence 1394 AA;
 Query Match 52.9%; Score 1337; DB 20; Length 1394;
 Best Local Similarity 55.4%; Pred. No. 2.2e-95;
 Matches 200; Conservative 61; Mismatches 88; Indels 12; Gaps 2;
 QY 13 CLTEDEKAAARVDQENIRLLEQKKQDRGELKLLILGPGESGKSTFIKOWRIIHGAGYSE 72
 Db 1045 clseakearrindelqrdrdarrelklllgtsgskstfkgmrlhngsgysd 1104
 QY 73 EERKGRPLVYQNIIVSMRAMTEAMERLOIP:SRPESKHHASLVMSQDPKYKVTTFEKRYA 132
 Db 1105 edkrgftklvyqnlftamqamramdtkltpkyehnkahaqvlrvdevkvsafenpyv 1164
 QY 133 AAMQWLWDGIRACRYERREFFHLLDSAVYVLSHLERITEGYVPTAQDVLSRMPPTGI 192
 Db 1165 daikslwvdlqecydrreyqlsdsktyllndlradvapylptqdvrvrvptgti 1224
 QY 193 NEYCFVSQVNTLRIVDVGQSKSRKWHCFENVALIYASLSEYDQCLEENQENRMK 252
 Db 1225 leypfdlqsvifrmvvggrsrrkwhcfenvtslmflvalseydvqlvivesdnenme 1284
 QY 253 ESALFPGTILELPWFKSTSVILEFNKTDILEEKIPTSHLATYPPSFGPKQDAEAAKRFI 312
 Db 1285 eskalftrltiypwfnqssvllfnkdlleeklmshlyvdyfpeydgprdaqarefi 1344
 QY 313 LDMYTRMTGCVDPGSGKSGKARSRLFSHYTCATDTQNIKRVKDVRSVLAARYLDEIN 372
 Db 1345 lkmfvd-----npdsdk-----liyshtcatdenirfvfaavkdtlqnlkdcg 1392
 QY 373 L 373
 Db 1393 L 1393
 RESULT 9
 AAY49134
 ID AAY49134 standard; Protein: 1397 AA.
 AC AAY49134;
 XX
 XX 07-JAN-2000 (first entry)
 DE pmGluR2/Car*Galpha15+3A1a linker fusion construct protein sequence.
 XX G-protein fusion receptor; Car; calcium receptor; GluR; head injury;
 KW metabotropic glutamate receptor; GABABR; chimeric receptor; stroke;
 KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
 KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
 KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
 KW cognitive disorder.
 XX
 OS Homo sapiens.
 XX WO9951641-A1.
 PN 14-OCT-1999.
 PD
 PF 02-APR-1999; 99WO-US07333.
 XX 03-APR-1998; 98US-0080671.
 PR

QY 73 EERKGRPLVYQNIYFVSMRAMTEAMERLQIPSPRESKHASLVMSODPYKVTTFEKRYA 132
 Db 1129 edkrgftklvygnifcamqamramdtlkipykayehnkahaglvrevdvkvsafepv 1188
 QY 133 AAMQWLWDAGIRACYSERREFFHLLDSAVYLSHLERITEEGYVPTAQDLRSRMTTGI 192
 Db 1189 daikslwvpgiqecydrreyqldstkyindlrdvdpaylptqgdlrvvptgti 1248
 QY 193 NEYCSVQKTNLRIVDVGOKSERKKWHCFENVTALIVLASLSEYDQCLEENNOENRMK 252
 Db 1249 leypfdlqsvifrmvvgqqrserkwhcfenvtalsimflvalseydqvivesdnenrme 1308
 QY 253 ESALFQITILELPWFKSTSVILFNKTDILEKIPTSHLATYFPFQGPQDAEAAKRFI 312
 Db 1309 eskalftrtiitypwnsqsvillfinkdlleeklmshlvdyfpeydgqrdagaarefi 1368
 QY 313 LDMYTRMYTCVDPGEGSKKARSRLFSHYTCATDTONIRKVFKDRSDSVLARYLDEIN 372
 Db 1369 lkmfvdl-----npdsdk-----liyshtcatdenirfvfaavkdtlqlnkdcg 1416
 QY 373 L 373
 Db 1417 L 1417

RESULT 11

AA49133
 ID AAY49133 standard; Protein; 1323 AA.

AC AAY49133;

DT 07-JAN-2000 (first entry)

DE GABA-BR1a*Gqo5 fusion construct protein sequence.

KW G-protein fusion receptor; CaR; calcium receptor; GluR; head injury;
 KW metabotropic glutamate receptor; GABAB; chimeric receptor; stroke;
 KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
 KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
 KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
 KW cognitive disorder.

OS Homo sapiens.

XX WO9951641-A1.

XX 14-OCT-1999.

XX 02-APR-1999; 99WO-US07333.

XX 03-APR-1998; 98US-0080671.

XX (NPSp-) NPS PHARM INC.

XX Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
 PI Simin RT;

XX WPI: 1999-610995/52.

XX DR N-P5DB; AAZ31064.

XX New G-protein fusion receptors and chimeras containing domains from
 XX different receptors, used to screen for modulators, potentially useful
 XX e.g. for treating or preventing stroke or Alzheimer's disease -
 XX Disclosure; Fig 14; 255pp; English.

XX The invention relates to G-protein fusion receptors (1) comprising:
 XX (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and
 XX intracellular (ICD) domains, each chosen independently from a CaR
 XX (calcium receptor), GluR (metabotropic glutamate receptor) and GABABR
 XX (gamma-aminobutyric acid receptor); (2) an optional linker attached to
 XX the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the

CC linker. (1), and recombinant chimeric receptors (CR) without the GP
 CC component are used to assess function of the various domains and to
 CC identify compounds (e.g. allosteric modulators or antagonists) that act
 CC on these domains. The modulators are potentially useful for treating or
 CC preventing diseases associated with the receptors, e.g. stroke, head or
 CC spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anoxia,
 CC Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive
 CC disorders and depression. Nucleic acid (II) that encodes (I) is used:
 CC (1) for recombinant production of corresponding proteins; and (2) to
 CC produce cells used in screening of modulators. Use of CaR and mGluR
 CC domains allows presentation of GABAB domains, to a binding agent, in a
 CC form more like the natural domain structure compared with use of
 CC incomplete receptors lacking one or more domains. By shuffling different
 CC domains, agents can be identified that affect particular domains of a
 CC receptor.

XX Sequence 1323 AA;

Query Match 52.8%; Score 1035.5; DB 20; Length 1323;
 Best Local Similarity 55.4%; Pred No. 2.9e-95;
 Matches 200; Conservative 62; Mismatches 82; Indels 17; Gaps 3;

QY 13 CLTEDEKAAARVDOEINRILLLEOKKODRGELKLLLLGPGESGKSTFIKOMRIHAGYSE 72
 Db 974 clseeakearrindeierqlrrdkdrareklllllgsgskcfikqmrllhngsyd 1033
 QY 73 EERKGRPLVYQNIYFVSMRAMTEAMERLQIPSPRESKHASLVMSODPYKVTTFEKRYA 132
 Db 1034 edkrgftklvygnifcamqamramdtlkipykayehnkahaglvrevdvkvsafepv 1093
 QY 133 AAMQWLWDAGIRACYSERREFFHLLDSAVYLSHLERITEEGYVPTAQDLRSRMTTGI 192
 Db 1094 daikslwvpgiqecydrreyqldstkyindlrdvdpaylptqgdlrvvptgti 1153
 QY 193 NEYCSVQKTNLRIVDVGOKSERKKWHCFENVTALIVLASLSEYDQCLEENNOENRMK 252
 Db 1154 leypfdlqsvifrmvvgqqrserkwhcfenvtalsimflvalseydqvivesdnenrme 1213
 QY 253 ESALFQITILELPWFKSTSVILFNKTDILEKIPTSHLATYFPFQGPQDAEAAKRFI 312
 Db 1214 eskalftrtiitypwnsqsvillfinkdlleeklmshlvdyfpeydgqrdagaarefi 1273
 QY 313 LDMYTRMYTCVDPGEGSKKARSRLFSHYTCATDTONIRKVFKDRSDSVLARYLDEIN 372
 Db 1274 lkmfvdl-----npdsdk-----liyshtcatdenirfvfaavkdtl-----qln 1316
 QY 373 L 373
 Db 1317 L 1317

RESULT 12

AA49132
 ID AAY49132 standard; Protein; 1303 AA.

AC AAY49132;

DT 07-JAN-2000 (first entry)

DE GABA-BR2*Gqo5 fusion construct protein sequence.

KW G-protein fusion receptor; CaR; calcium receptor; GluR; head injury;
 KW metabotropic glutamate receptor; GABAB; chimeric receptor; stroke;
 KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
 KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
 KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
 KW cognitive disorder.

OS Homo sapiens.

XX WO9951641-A1.

PD 14-OCT-1999.

XX 02-APR-1999; 99WO-US07333.

XX 03-APR-1998; 98US-0080671.

XX (NPSF-) NPS PHARM INC.

XX Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;

PI Simin RT;

XX WPI; 1999-610995/52.

DR N-PSDB; AA231063.

XX New G-protein fusion receptors and chimeras containing domains from

PT different receptors, used to screen for modulators, potentially useful

PT e.g. for treating or preventing stroke or Alzheimer's disease

XX Disclosure; Fig 14; 255pp; English.

XX The invention relates to G-protein fusion receptors (I) comprising:

CC (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and

CC intracellular (ICD) domains, each chosen independently from a car

CC (calcium receptor), GluR (metabotropic glutamate receptor) and GABABR

CC (gamma-aminobutyric acid receptor); (2) an optional linker attached to

CC the C-terminus of ICD; and (3) a G-protein (Gp) linked to ICD or the

CC linker. (I), and recombinant chimeric receptors (CR) without the Gp

CC component, are used to assess function of the various domains and to

CC identify compounds (e.g. allosteric modulators or antagonists) that act

CC on these domains. The modulators are potentially useful for treating or

CC preventing diseases associated with the receptors, e.g. stroke, head or

CC spinal cord injury, epilepsy, ischaemia, hypoglycaemia, cognitive

CC Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive

CC disorders and depression. Nucleic acid (II) that encodes (I) is used:

CC (1) for recombinant production of corresponding proteins; and (2) to

CC produce cells used in screening for modulators. Use of Car and mGluR

CC domains allows presentation of GABABR domains, to a binding agent, in a

CC form more like the natural domain structure compared with use of

CC incomplete receptors, lacking one or more domains. By shuffling different

CC domains, agents can be identified that affect particular domains of a

CC receptor.

XX

SQ Sequence 1303 AA;

QY 373 L 373

DB 1297 L 1297

RESULT 13

AA52705

ID AAY52705 standard; Protein; 359 AA.

XX AAY52705;

XX 26-JAN-2000 (first entry)

XX Human G-alpha-11 protein.

XX Identification; genetic target; gene modulation; human;

XX anticense oligonucleotide; phosphorothioate; target validation;

XX nucleotide sequence-based technology; antisense drug discovery.

XX Homo sapiens.

XX WO9553101-A1.

XX 21-OCT-1999.

XX 13-APR-1999; 99WO-US08268.

XX 13-APR-1998; 98US-0081483.

XX 28-APR-1998; 98US-0067638.

XX (ISIS-) ISIS PHARM INC.

XX Cowser LM, Baker BF, McNeil J, Freier SM, Sasmor HM, Brooks DG;

XX Ohasi C, Wyatt JR, Borchers AH, Vickers TA;

XX WPI; 1999-620446/53.

XX N-PSDB; AA241093.

XX Identifying compounds which modulate expression of nucleic acids, used

XX to provide compounds having defined physical, chemical or bioactive

XX properties, e.g. antisense activity

XX Example 26; Page 226-228; 264pp; English.

XX A method has been developed of defining a set of compounds that modulate

XX the expression of a target nucleic acid (TNA) sequence via binding of

XX the compounds with the TNA sequence. The method comprises generating a

XX library of virtual compounds in silico according to defined criteria,

XX and evaluating in silico the binding of the virtual compounds with the

XX TNA according to defined criteria. Also described are: (1) a method of

XX defining a set of oligonucleotides (ONS) that modulate the expression of

XX a TNA sequence via binding of the ONS with the TNA sequence comprising

XX generating a library of virtual compounds in silico according to defined

XX criteria, and evaluating in silico the binding of the virtual ONS with

XX the TNA according to defined criteria; and (2) a method of defining a

XX set of compounds that modulate the expression of a TNA sequence via

XX binding of the compounds with the TNA. The methods can be used for the

XX generation and identification of synthetic compounds having defined

XX physical, chemical or bioactive properties. Information gathered from

XX assays of such compounds is used to identify nucleic acid sequences that

XX are tractable to a variety of nucleotide sequence-based technologies,

XX e.g. antisense drug discovery and target validation. AA240852 to

XX AA241220, and AAY52701 to AAY52706, represent sequences used in the

XX exemplification of the present invention.

XX Sequence 359 AA;

QY 13 CLTEDEKAAARVDQEIINRILLEGKQKQDRGELKLLLLGPGESGKSTFIKQMRHHCAGYSE 72

Query Match 52.7%; Score 1033; DB 20; Length 359;

Best Local Similarity 55.5%; Pred. No. 7.2e-96;

Matches 201; Conservative 65; Mismatches 84; Indels 12; Gaps 2;

```

Db 10 clsdvkeskrinaeiekqlrrdrrelkllllgtgesgskstfrikqmriihgagys 69
Qy 73 EERKGRPLVYQNIYFVSMRAMIEMERLOTPFSPRSKHHASLVMSQDPYKVTTEKRYA 132
Db 70 edkrgtklvyqnlftamqamirametkilykyeqkanallirevdekvttefhqy 129
Qy 133 AAMQWLWRDAGIRACYERREPHLLDSAVYVLSHLERITEEGYVPTAOVLRSMPTTGI 192
Db 130 saiktledpgiqecydrreyqlsdsakylldvdratiglyptqdvrvrvptgtl 189
Qy 193 NEYCFVQKTNLRIVDVGGQSKSRKWHCFENVIALIYLASEYDQCLEENNQENRMK 252
Db 190 iepyfdalenlfrmdvvgggsrfrkwhcfenvtsimflvalseydqvlivesdnenme 249
Qy 253 ESALFEGTILELPWPKSTSVILFLNKTDILEKIPKSHLATYPPSFGPKODAAKRFI 312
Db 250 eskalftrltitypwnssvllflnkddilekilyshlvdyfpefdgpgardqaarefi 309
Qy 313 LDMYTRMTGCVDPGSGKSGARSRLFSHYTCATDTONIRKVKDVRDVLARYLDEIN 372
Db 310 pkmfvdj-----npsdk-----liyshfcatdenirfvaavkdtllqlnkeyn 357
Qy 373 LL 374
Db 358 LV 359

```

```

RESULT 14
AAV29789
ID AAV29789 standard; Protein; 359 AA.
XX AC AAV29789;
XX DT 15-NOV-1999 (first entry)
XX DE Human G-alpha-11 protein sequence.
XX KW Human; G-alpha-11; antisense oligonucleotide; inhibition; expression;
XX KW phosphorothioate.
XX OS Homo sapiens.
XX PN US5951455-A.
XX PD 14-SEP-1999.
XX PF 04-DEC-1998; 98US-0205922.
XX PR 04-DEC-1998; 98US-0205922.
XX PA (ISIS-) ISIS PHARM INC.
XX PI Cowsert LM;
XX DR WPI; 1999-539140/45.
XX DR N-PSDB; AAZ19461.
XX PT Inhibitory antisense compounds useful for the treatment of diseases
XX PS associated with G-alpha-11
XX Example 13; Column 45-48; 38pp; English.
XX The present invention describes inhibitory antisense compounds of 8-30
XX nucleotides, targeted to a nucleic acid molecule encoding human
XX G-alpha-11. The present sequence represents human G-alpha-11. AAZ19468
XX to AAZ19547 represent human G-alpha-11 phosphorothioate antisense
XX oligonucleotides given in the present invention. The oligonucleotides
XX may be useful for the treatment of diseases associated with G-alpha-11.
XX Sequence 359 AA;

```

```

Query Match 52.7%; Score 1333; DB 20; Length 359;
Best Local Similarity 55.5%; Pred. No. 7.2e-96; Indels 12; Caps 2;
Matches 201; Conservative 65; Mismatches 84;
Qy 13 CLTEDEKAAARVOENRILLEOKKQDRGELKLLILGPGESGKSTFIKQMRIIHGAGYSE 72
Db 10 clsdvkeskrinaeiekqlrrdrrelkllllgtgesgskstfrikqmriihgagys 69
Qy 73 EERKGRPLVYQNIYFVSMRAMIEMERLOTPFSPRSKHHASLVMSQDPYKVTTEKRYA 132
Db 70 edkrgtklvyqnlftamqamirametkilykyeqkanallirevdekvttefhqy 129
Qy 133 AAMQWLWRDAGIRACYERREPHLLDSAVYVLSHLERITEEGYVPTAOVLRSMPTTGI 192
Db 130 saiktledpgiqecydrreyqlsdsakylldvdratiglyptqdvrvrvptgtl 189
Qy 193 NEYCFVQKTNLRIVDVGGQSKSRKWHCFENVIALIYLASEYDQCLEENNQENRMK 252
Db 190 iepyfdalenlfrmdvvgggsrfrkwhcfenvtsimflvalseydqvlivesdnenme 249
Qy 253 ESALFEGTILELPWPKSTSVILFLNKTDILEKIPKSHLATYPPSFGPKODAAKRFI 312
Db 250 eskalftrltitypwnssvllflnkddilekilyshlvdyfpefdgpgardqaarefi 309
Qy 313 LDMYTRMTGCVDPGSGKSGARSRLFSHYTCATDTONIRKVKDVRDVLARYLDEIN 372
Db 310 pkmfvdj-----npsdk-----liyshfcatdenirfvaavkdtllqlnkeyn 357
Qy 373 LL 374
Db 358 LV 359

```

```

RESULT 15
AAV49125
ID AAV49125 standard; Protein; 359 AA.
XX AC AAV49125;
XX DT 07-JAN-2000 (first entry)
XX DE Chimeric Gq15 protein sequence.
XX KW G-protein fusion receptor; CaR; calcum receptor; GluR; head injury;
XX KW metabotropic glutamate receptor; GABAAR; chimeric receptor; stroke;
XX KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
XX KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
XX KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
XX KW cognitive disorder.
XX OS Homo sapiens.
XX PA WO9951641-A1.
XX PI 14-OCT-1999.
XX PF 02-APR-1999; 99WO-US07333.
XX PR 03-APR-1998; 98US-0080671.
XX PA (NPS- ) NPS PHARM INC.
XX PI Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
XX PI Simlin RT;
XX DR WPI; 1999-610995/52.
XX DR N-PSDB; AAZ31056.
XX PT New G-protein fusion receptors and chimeras containing domains from
XX PT different receptors, used to screen for modulators, potentially useful
XX PT e.g. for treating or preventing stroke or Alzheimer's disease
XX PS Disclosure; Page 196-197; 255pp; English.

```

Search completed: September 6, 2001, 10:54:33
Job time: 53 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 6, 2001, 10:53:40 ; Search time 13.15 Seconds
(without alignments)
974.262 Million cell updates/sec

Title: US-08-878-801-2

Perfect score: 1960
Sequence: 1 MARSLTRCCPWCLTEDEKA.....VFKDVRDSVLARYLDEINLL 374

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1960	100.0	374	1 GB15_HUMAN	P30679 homo sapien
2	1722	87.9	374	1 GB15_RAT	O88302 rattus norv
3	1705	87.0	374	1 GB15_MOUSE	P30678 mus musculu
4	1052	53.7	353	1 GBQ_XENLA	P38410 xenopus lae
5	1047	53.4	353	1 GBQ_CANFA	Q28294 canis fami
6	1046	53.4	353	1 GBQ_RAT	P82471 rattus norv
7	1046	53.4	359	1 GB11_BOVIN	P38409 bos taurus
8	1044	53.3	353	1 GBQ_MOUSE	P21279 mus musculu
9	1042	53.2	353	1 GBQ_HUMAN	P50148 homo sapien
10	1040	53.1	359	1 GB11_MOUSE	P29992 homo sapien
11	1038	53.0	359	1 GB11_MOUSE	P21278 mus musculu
12	1036	52.9	353	1 GBQ_HOMAM	P19150 homarus ame
13	1035	52.8	359	1 GB11_MELGA	P45645 melagris g
14	1030	52.6	359	1 GB11_RAT	Q9J102 rattus norv
15	1026.5	52.4	359	1 GBQ_XENLA	P43444 xenopus lae
16	1026	52.3	353	1 GBQ1_DROME	P23625 drosophila
17	1026	52.3	353	1 GBQ3_DROME	P54400 drosophila
18	1017.5	51.9	354	1 GBQ_LOLFO	P38412 loligo forb
19	1013	51.7	354	1 GB14_XENLA	O73819 xenopus lae
20	1013	51.7	355	1 GB14_BOVIN	P38408 bos taurus
21	1013	51.7	355	1 GB14_MOUSE	P30677 mus musculu
22	1005	51.3	355	1 GB14_HUMAN	O95837 homo sapien
23	995	50.8	353	1 GBQ_LYMST	P38411 lymnaea sta
24	989	50.5	353	1 GBQ_PATYE	O15975 patinopecte
25	787.5	40.2	352	1 GBAL_EMENI	Q00743 emericella
26	766.5	39.1	353	1 GBAL_NEUCR	Q05425 neurospora
27	764	39.0	352	1 GBAL_COCH	O74227 cochliobol
28	762	38.9	354	1 GB02_DROME	P16377 drosophila
29	760.5	38.8	354	1 GB01_DROME	P16378 drosophila
30	760	38.8	352	1 GBAL_COPCO	P30675 coprinus co
31	759.5	38.8	352	1 GBAL_CRYPA	Q00580 cryphonectr
32	759.5	38.8	353	1 GBAL_USTMA	P87032 ustilago ma
33	757.5	38.6	352	1 GBAL_COLTR	O42784 colletotric

RESULT 1

GB15_HUMAN	374 AA
ID	GB15_HUMAN STANDARD; PRT; 374 AA
AC	P30679; O75247;
DT	01-APR-1993 (Rel. 25, Created)
DT	01-APR-1993 (Rel. 25, Last sequence update)
DT	01-OCT-2000 (Rel. 40, Last annotation update)
DE	GUANINE NUCLEOTIDE-BINDING PROTEIN, ALPHA-15 SUBUNIT (ALPHA-16).
GN	GNA15 OR GNA16.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RP	SEQUENCE FROM N.A.
RX	MEDLINE=91288509; PubMed=1905813;
RA	Amatruda T.T. III, Steele D.A., Slepak V.Z., Simon M.I.;
RT	"G alpha 16, a G protein alpha subunit specifically expressed in
RL	hematopoietic cells.";
RL	Proc. Natl. Acad. Sci. U.S.A. 88:5587-5591(1991).
RP	SEQUENCE FROM N.A.
RA	Lamerdin J.E., McCreedy P.M., Skowronski E., Adamson A.W.,
RA	Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stillwagen S.,
RA	Phan H., Velasco N., Do L., Regala W., Terry A., Ganes J.,
RA	Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA	Liu S., Attix C., Andreise T., Frankheim M., Amico-Keller G.,
RA	Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA	Kronmiller B., Arellano A., Montgomery M., Ow D., Nolan M., Trong S.,
RA	Kobayashi A., Olsen A.S., Carrano A.V.;
RL	Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC	-!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC	INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC	SIGNALING SYSTEMS.
CC	-!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC	THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC	-!- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN HEMATOPOIETIC CELLS.
CC	-!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G10).
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; M63904; AAA35860.1;
DR	EMBL; AC005264; AAC25612.1;
DR	EMBL; AC005262; AAC25616.1;
DR	PIR; A41096; A41096.
DR	HSP; P04896; IAZT.
DR	MIM; 139314;
DR	InterPro; IPR000654;
DR	InterPro; IPR001019;
DR	Pfam; PF00503; G-alpha; 1.

ALIGNMENTS

P51876 helisoma tr
P30682 lymnaea sta
P27601 mus musculu
O74259 sporothrix
P27044 xenopus lae
O13315 magnaporthe
P30683 lymnaea sta
P50146 gallus gall
P51877 helisoma tr
P38401 cavia porce
P87383 oryzias lat
P29348 rattus norv

34 756 38.6 353 1 GB1_HELTI
35 755 38.5 353 1 GB1_LYMST
36 755 38.5 377 1 GB13_MOUSE
37 753 38.4 352 1 GBAL_SPOSC
38 750.5 38.3 353 1 GB11_XENLA
39 749.5 38.2 352 1 GBAL_WAGR
40 748.5 38.2 353 1 GB0_LYMST
41 748.5 38.2 353 1 GB11_CHICK
42 747.5 38.1 353 1 GB0_HELTI
43 746.5 38.1 353 1 GB11_CAVPO
44 746.5 38.1 353 1 GB11_ORYLA
45 746.5 38.1 353 1 GB13_RAT

DR PRINTS; PRO0318; GPROTEINA.
 DR PRINTS; PRO0442; GPROTEINAQ.
 KW GTP-binding; Transducer; Multigene family; ADP-ribosylation.
 FT NP_BIND 49 56 GTP (BY SIMILARITY).
 FT NP_BIND 208 212 GTP (BY SIMILARITY).
 FT NP_BIND 277 280 GTP (BY SIMILARITY).
 FT NP_BIND 186 186 ADP-RIBOSYL[1] (BY ACTION OF CTX)
 FT CONFLICT 147 147 C -> Y (IN REF. 2).
 SQ SEQUENCE 374 AA; 43508 MW; 8127AC16FC212507 CRC64;

Query Match 100.0%; Score 1960; DB 1; Length 374;
 Best Local Similarity 100.0%; Pred No. 2.9e-150;
 Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARSLTRCCPWCCLTDEKAAARVDQEIINRILLEQKKQDRGELKLLLLGPGESGKSTFIK 60
 DB 1 MARSLTRCCPWCCLTDEKAAARVDQEIINRILLEQKKQDRGELKLLLLGPGESGKSTFIK 60
 QY 61 OMRIIHGAGYSEERGFRLVYQNFVSMRAMIEMERLQIPFSPKSHASLVMSQD 120
 DB 61 OMRIIHGAGYSEERGFRLVYQNFVSMRAMIEMERLQIPFSPKSHASLVMSQD 120
 QY 121 PKVTTFFERYAAAMQWLMDAGIRACRYERRERFHLDSAVYLYSLHLRITEREYVPTAQ 180
 DB 121 PKVTTFFERYAAAMQWLMDAGIRACRYERRERFHLDSAVYLYSLHLRITEREYVPTAQ 180
 QY 121 PKVTTFFERYAAAMQWLMDAGIRACRYERRERFHLDSAVYLYSLHLRITEREYVPTAQ 180
 DB 121 PKVTTFFERYAAAMQWLMDAGIRACRYERRERFHLDSAVYLYSLHLRITEREYVPTAQ 180
 QY 181 DVLSRMPPTTGINEYCFVQKTNLRIVDVGQKSERKKWIHCNFENVIALIYLASLSEYDQ 240
 DB 181 DVLSRMPPTTGINEYCFVQKTNLRIVDVGQKSERKKWIHCNFENVIALIYLASLSEYDQ 240
 QY 241 CLEENQENRMKESLALFGLTILELPWFKSTSVILFLNKTDLLEEKIPTSHLATYFPFSG 300
 DB 241 CLEENQENRMKESLALFGLTILELPWFKSTSVILFLNKTDLLEEKIPTSHLATYFPFSG 300
 QY 301 PKQDAEAAKRFILDMYTRMYTCVGDGPEGSKKGGARSRLFSHYTCATDTQNTKRVKDV 360
 DB 301 PKQDAEAAKRFILDMYTRMYTCVGDGPEGSKKGGARSRLFSHYTCATDTQNTKRVKDV 360
 QY 361 DSVLARYLDEINLL 374
 DB 361 DSVLARYLDEINLL 374

RESULT 2
 GB15_RAT
 ID GB15_RAT STANDARD; PRT: 374 AA.
 AC GB15_RAT
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GUANINE NUCLEOTIDE-BINDING PROTEIN, ALPHA-15 SUBUNIT.
 GN GNA15.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98351893; PubMed=9685675;
 RA Kusakabe Y., Yamaguchi E., Tanemura K., Kameyama K., Chiba N.,
 RA Arai S., Emori Y., Abe K.,
 RT "Identification of two alpha-subunit species of GTP-binding proteins,
 RT G alpha 15 and G alpha 4, expressed in rat taste buds.";
 RL Biochim. Biophys. Acta 1403:265-272(1998).
 CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
 CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
 CC SIGNALING SYSTEMS.
 CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
 CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
 CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(O)).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AB015308; BAA28927.1; -
 DR HSP; P04896; JAZT.
 DR InterPro; IPR000634; -
 DR InterPro; IPR001019; -
 DR Pfam; PF00503; G-alpha; 1.
 DR PRINTS; PRO0318; GPROTEINA.
 DR PRINTS; PRO0442; GPROTEINAQ.
 KW GTP-binding; Transducer; Multigene family; ADP-ribosylation.
 FT NP_BIND 49 56 GTP (BY SIMILARITY).
 FT NP_BIND 208 212 GTP (BY SIMILARITY).
 FT NP_BIND 277 280 GTP (BY SIMILARITY).
 FT MOD_RES 186 186 ADP-RIBOSYL[1] (BY ACTION OF CTX)
 FT (BY SIMILARITY).
 SQ SEQUENCE 374 AA; 43331 MW; E5E65D32D41EC0EB CRC64;

Query Match 87.9%; Score 1122; DB 1; Length 374;
 Best Local Similarity 86.1%; Pred No. 3.8e-131;
 Matches 322; Conservative 30; Mismatches 22; Indels 0; Gaps 0;

QY 1 MARSLTRCCPWCCLTDEKAAARVDQEIINRILLEQKKQDRGELKLLLLGPGESGKSTFIK 60
 DB 1 MARSLTRCCPWCCLTDEKAAARVDQEIINRILLEQKKQDRGELKLLLLGPGESGKSTFIK 60
 QY 61 OMRIIHGAGYSEERGFRLVYQNFVSMRAMIEMERLQIPFSPKSHASLVMSQD 120
 DB 61 OMRIIHGAGYSEERGFRLVYQNFVSMRAMIEMERLQIPFSPKSHASLVMSQD 120
 QY 121 PKVTTFFERYAAAMQWLMDAGIRACRYERRERFHLDSAVYLYSLHLRITEREYVPTAQ 180
 DB 121 PKVTTFFERYAAAMQWLMDAGIRACRYERRERFHLDSAVYLYSLHLRITEREYVPTAQ 180
 QY 181 DVLSRMPPTTGINEYCFVQKTNLRIVDVGQKSERKKWIHCNFENVIALIYLASLSEYDQ 240
 DB 181 DVLSRMPPTTGINEYCFVQKTNLRIVDVGQKSERKKWIHCNFENVIALIYLASLSEYDQ 240
 QY 241 CLEENQENRMKESLALFGLTILELPWFKSTSVILFLNKTDLLEEKIPTSHLATYFPFSG 300
 DB 241 CLEENQENRMKESLALFGLTILELPWFKSTSVILFLNKTDLLEEKIPTSHLATYFPFSG 300
 QY 301 PKQDAEAAKRFILDMYTRMYTCVGDGPEGSKKGGARSRLFSHYTCATDTQNTKRVKDV 360
 DB 301 PKQDAEAAKRFILDMYTRMYTCVGDGPEGSKKGGARSRLFSHYTCATDTQNTKRVKDV 360
 QY 361 DSVLARYLDEINLL 374
 DB 361 DSVLARYLDEINLL 374

RESULT 3
 GB15_MOUSE
 ID GB15_MOUSE STANDARD; PRT: 374 AA.
 AC P30678;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GUANINE NUCLEOTIDE-BINDING PROTEIN, ALPHA-15 SUBUNIT.
 GN GNA15 OR GNA-15.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92052208; PubMed=1946421;

[illegible]

```

FT  MOD_RES 177 177 ADP-RIBOSYL[1] (BY ACTION OF CTX)
FT  CONFLICT 10 10 E -> K (IN REF. 2).
FT  CONFLICT 79 79 S -> T (IN REF. 2).
FT  CONFLICT 158 158 L -> V (IN REF. 2).
FT  CONFLICT 164 164 H -> Q (IN REF. 2).
SQ  SEQUENCE 353 AA; 41541 MW; 2310142412184553 CRC64;

Query Match 53.7%; Score 1052; DB 1; Length 353;
Best Local Similarity 56.4%; Pred. No. 2.3e-77;
Matches 204; Conservative 60; Mismatches 86; Indels 12; Gaps 2;

QY 13 CLTEDEKAAARVDQENRILLEOKKODRGELKLLLLGPGESGKSTFTKQRIIHGAGYSE 72
Db 4 CLSEEAEEARRINDEIERLRDRKRDARRELKLLLLGTGESGKSTFTKQRIIHGAGYSD 63
QY 73 EERKGRPLVYQNIYFVSMRAMIEMERLQIPSPRPESKHHASLVMSQDPKYVTFEKRYA 132
Db 64 EDKRGFTKLVYQNIYFSAQMIAMETLKIPIYEHKHAHQAHLVREVDVEKVSFENPYV 123
QY 133 AMQWLWRDAGIRACVRRERFHLDSAVYYLSHLERITEEGYVPTAQDVLRSMPPTGI 192
Db 124 DAIKYLWDPGIGQCYDRRREYQLSYSTKYLLNDLRVADPAYLPYQDQVLRVPTTGI 183
QY 193 NEYCFVQKTNLRIVDVGQKSRKKWIIHCFENVIALIYLSLSEYDOCLEENNQENRMK 252
Db 184 IEYFQLQSVIFRMVDVGQSGRRKWIHCFENVTSIMFVLSEYDQVLVESDNENRME 243
QY 253 ESALFGLTILELPWFKSTSVILFNKTDILEEKIPTSHLATYPPSQGPKQDAEAAKRFI 312
Db 244 ESKALFRTIITYPWFQNSVILFNKKDLLEIKIMYSHLVDFYPEYDGPQDAQAAREFI 303
QY 313 LDMYTRMTGCVDPGPGSKGARSRLFSHYTCATDTQNIKRVKDVRSVLARYLDEIN 372
Db 304 LKMEVDL-----NPDSK-----IYSHFTCATDTENIRFVFAVKDTILQNLKEYN 351
QY 373 LL 374
Db 352 LV 353

RESULT 5
GBQ_CANFA
ID GBQ_CANFA STANDARD; PRT; 353 AA.
AC Q28294;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE GUANINE NUCLEOTIDE-BINDING PROTEIN G(O), ALPHA SUBUNIT.
GN GNAO.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96433124; PubMed=8836152;
RA Johnson G.J., Leis L.A., Dunlop P.C.;
RT "Specificity of G alpha q and G alpha 11 gene expression in platelets
RT and erythrocytes. Expressions of cellular differentiation and species
RT differences."
RL Biochem. J. 318:1023-1031(1996).
CC -1- INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS.
CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(O)).
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L76257; AAB05548.1; ALT_INIT.
DR HSSP; P04896; IAZT.
DR InterPro; IPR000654; -.
DR Pfam; PF00503; G-alpha.1.
DR PRINTS; PR00318; GPROTEINA.
DR PRINTS; PR00442; GPROTEINAO.
KW GTP-binding; Transducer; Multigene family; ADP-ribosylation;
KW Palmitate; Lipoprotein.
FT LIPID 3 3 PALMITATE (BY SIMILARITY).
FT LIPID 4 4 GTP (BY SIMILARITY).
FT NP_BIND 40 47 GTP (BY SIMILARITY).
FT NP_BIND 199 203 GTP (BY SIMILARITY).
FT NP_BIND 268 271 ADP-RIBOSYL[1] (BY ACTION OF CTX)
FT MOD_RES 177 177 (BY SIMILARITY).
SQ SEQUENCE 353 AA; 41467 MW; E7737307B1F4904C CRC64;

Query Match 53.4%; Score 1047; DB 1; Length 353;
Best Local Similarity 55.8%; Pred. No. 5.7e-77;
Matches 202; Conservative 61; Mismatches 87; Indels 12; Gaps 2;

QY 13 CLTEDEKAAARVDQENRILLEOKKODRGELKLLLLGPGESGKSTFTKQRIIHGAGYSE 72
Db 4 CLSEEAEEARRINDEIERLRDRKRDARRELKLLLLGTGESGKSTFTKQRIIHGAGYSD 63
QY 73 EERKGRPLVYQNIYFVSMRAMIEMERLQIPSPRPESKHHASLVMSQDPKYVTFEKRYA 132
Db 64 EDKRGFTKLVYQNIYFSAQMIAMETLKIPIYEHKHAHQAHLVREVDVEKVSFENPYV 123
QY 133 AMQWLWRDAGIRACVRRERFHLDSAVYYLSHLERITEEGYVPTAQDVLRSMPPTGI 192
Db 124 DAIKYLWDPGIGQCYDRRREYQLSYSTKYLLNDLRVADPAYLPYQDQVLRVPTTGI 183
QY 193 NEYCFVQKTNLRIVDVGQKSRKKWIIHCFENVIALIYLSLSEYDOCLEENNQENRMK 252
Db 184 IEYFQLQSVIFRMVDVGQSGRRKWIHCFENVTSIMFVLSEYDQVLVESDNENRME 243
QY 253 ESALFGLTILELPWFKSTSVILFNKTDILEEKIPTSHLATYPPSQGPKQDAEAAKRFI 312
Db 244 ESKALFRTIITYPWFQNSVILFNKKDLLEIKIMYSHLVDFYPEYDGPQDAQAAREFI 303
QY 313 LDMYTRMTGCVDPGPGSKGARSRLFSHYTCATDTQNIKRVKDVRSVLARYLDEIN 372
Db 304 LKMEVDL-----NPDSK-----IYSHFTCATDTENIRFVFAVKDTILQNLKEYN 351
QY 373 LL 374
Db 352 LV 353

RESULT 6
GBQ_RAT
ID GBQ_RAT STANDARD; PRT; 353 AA.
AC P82471;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE GUANINE NUCLEOTIDE-BINDING PROTEIN G(O), ALPHA SUBUNIT.
GN GNAO.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RT TISSUE-Brain;

```


QY 133 AMOWLRDAGIRACVYERREFHLLDSAVYLSHLERITEEGYVPTAQDVLRSRMTPTGI 192
 Db 130 SAIKTLWDPGIEQYDRREYQLSDSAKYLLTDVDRATSGYLTQDQDLVRVPTPTGI 189
 QY 193 NEYCFVQKTNLRIVDVGOKSERKWIHCFFENFVIALIYLSLSEYDQCLFENNOENRMK 252
 Db 190 IEYFPDLNFIIFRMDVVGORSEKRWIHCFFENFVITSMFLVSEYDQCLFENNOENRMK 249
 QY 253 ESALFGLTELEPWFKSTSVILFNKTDILEEIKPTSLATYFPFSGPKODAAAKRFI 312
 Db 250 ESKALFTVITYPWFQNSVILFNKDLKDLKTLHSLVDYFPFSGPKODAAAKRFI 309
 QY 313 LDMYTRMTCVDPGSGKRSRLFSHYTCATQNTQIRKVKDVRDSVLARYLDEIN 372
 Db 310 LKMFVDL-----NPDSDK-----IIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYN 357
 QY 373 LL 374
 Db 358 LV 359

RESULT 8
 ID GBQ_MOUSE STANDARD; PRT: 353 AA.
 AC P21279;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GUANINE NUCLEOTIDE-BINDING PROTEIN G(O), ALPHA SUBUNIT.
 GN GNAO.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP TISSUE=Brain;
 RC TISSUE=Brain;
 RX MEDLINE=91067657; PubMed=2123549;
 RA Strathmann M., Simon M.I.;
 RT "G protein diversity: a distinct class of alpha subunits is present
 in vertebrates and invertebrates.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:9113-9117(1990).
 RN [2]
 RP PALMITOYLATION.
 RC TISSUE=Brain;
 RX MEDLINE=94043367; PubMed=8227083;
 RA Wedegaertner P.B., Chu D.H., Wilson P.T., Lewis M.J., Bourne H.R.;
 RT "Palmitoylation is required for signaling functions and membrane
 attachment of Gq alpha and Gs alpha.";
 RL J. Biol. Chem. 268:25001-25008(1993).
 CC -|- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
 INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
 SIGNALING SYSTEMS.
 CC -|- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
 CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
 CC -|- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(O)).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; M55412; AAA63306.1; ALT_INIT.
 DR EIR; A38414; RGM5Q.
 DR HSP; P04896; IAZT.
 DR MGD; MGI:95776; Gnaq.
 DR InterPro; IPR000654; -.
 DR InterPro; IPR001019; -.
 DR Pfam; PF00503; G-alpha; 1.
 DR PRINTS; PR003118; GPROTEINA.

DR PRINTS; PR00442; GPROTEINAQ.
 KW GTP-binding; Transducer; Multigene family; ADP-ribosylation;
 KW Palmitate; Lipoprotein.
 FT LIPID 3 PALMITATE.
 FT LIPID 4 4
 FT NP_BIND 40 47 GTP (BY SIMILARITY).
 FT NP_BIND 199 203 GTP (BY SIMILARITY).
 FT NP_BIND 268 271 GTP (BY SIMILARITY).
 FT MOD_RES 177 177 ADP-RIBOSYL[1] (BY ACTION OF CTX)
 FT (BY SIMILARITY).
 SQ SEQUENCE 353 AA; 41478 MW; C41B2AC11C674C5F CRC64;

Query Match 53.3%; Score 1344; DB 1; Length 353;
 Best Local Similarity 55.5%; Pred. No. 1e-76; 87; Indels 12; Gaps 2;
 Matches 201; Conservative 62; Mismatches 87;

QY 13 CLTEDEAAARVDOEINRILLOKQKODRGELKLLGLPGESGKSTFIKQRIIHGAGYSE 72
 Db 4 CLSEAEKARRINDEIERHVRDKRDARRELKLLGLTGESGKSTFIKQRIIHGAGYSD 63
 QY 73 EERKGRPLVYQNIIVSMRAMIEAMERLQIPSRPESKHHSALVMSODPYKVTTFEKRYA 132
 Db 64 EDKRGFTKLVIYQNIETAMQAMIRAMDTLKP/KYEHNKAHAQLVREYDVEKVSAPENPV 123
 QY 133 AMOWLRDAGIRACVYERREFHLLDSAVYLSHLERITEEGYVPTAQDVLRSRMTPTGI 192
 Db 124 DAISLWNPDPGIEQYDRREYQLSDSKYLNLDLRVADPSYLTQDQDLVRVPTPTGI 183
 QY 193 NEYCFVQKTNLRIVDVGOKSERKWIHCFFENFVIALIYLSLSEYDQCLFENNOENRMK 252
 Db 184 LEYFDLQSVIFRMDVVGORSEKRWIHCFFENFVITSMFLVSEYDQCLFENNOENRMK 243
 QY 253 ESALFGLTELEPWFKSTSVILFNKTDILEEIKPTSLATYFPFSGPKODAAAKRFI 312
 Db 244 ESKALFTVITYPWFQNSVILFNKDLKDLKTLHSLVDYFPFSGPKODAAAKRFI 303
 QY 313 LDMYTRMTCVDPGSGKRSRLFSHYTCATQNTQIRKVKDVRDSVLARYLDEIN 372
 Db 304 LKMFVDL-----NPDSDK-----IIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYN 351
 QY 373 LL 374
 Db 352 LV 353

RESULT 9
 ID GBQ_HUMAN STANDARD; PRT: 353 AA.
 AC P50148; Q13462; Q92471; O15108;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE GUANINE NUCLEOTIDE-BINDING PROTEIN G(O), ALPHA SUBUNIT.
 GN GNAO OR GAO.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96423032; PubMed=8825633;
 RA Dong Q., Shenker A., Way J., Haddad B.R., Lin K., Hughes M.R.,
 McBridge W.O., Spiegel A.M., Battey J.;
 RT "Molecular cloning of human G alpha q cDNA and chromosomal
 localization of the G alpha q gene (GNAO) and a processed
 pseudogene.";
 RL Genomics 30:470-475(1995).
 RN [2]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RC TISSUE=Prostate;
 RX MEDLINE=96256639; PubMed=8664309;
 RA Chen B., Leverette R.D., Schwinn D.A., Kwatra M.M.;

TISSUE-Hematopoietic;
 MEDLINE=96077138; PubMed=7492305;
 Thomas C.P., Dunn M.J., Mattera R.;
 Ca2+ signalling in K562 human erythroleukaemia cells: effect of
 dimethyl sulphoxide and role of G-proteins in thrombin- and
 thromboxane A2-activated pathways.;
 Biochem. J. 312:151-158(1995).
 -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
 INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
 SIGNALING SYSTEMS. ACTS AS AN ACTIVATOR OF PHOSPHOLIPASE C.
 -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
 THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
 -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch)

 EMBL: M69013; AAS58624.1;
 EMBL: AF011497; AAB64303.1;
 EMBL: AC005262; AAC25615.1;
 EMBL: L40630; AAA99949.1;
 PIR: A39394; RGHUY.
 HSSP: P04896; LAZT.
 MIM: 139313;
 InterPro: IPR000654;
 InterPro: IPR001019;
 Pfam: PF00503; G-alpha; 1.
 PRINTS: PR00318; GPROTEINA.
 PRINTS: PR00442; GPROTEINAQ.
 GTP-binding; transducer; Multigene family; ADP-ribosylation.
 NP_BIND 46 53 GTP (BY SIMILARITY).
 NP_BIND 205 209 GTP (BY SIMILARITY).
 NP_BIND 274 277 GTP (BY SIMILARITY).
 MOD_RES 183 183 ADP-RIBOSYL[1] (BY ACTION OF CTX)
 (BY SIMILARITY).
 CONFLICT 6 6 M -> I (IN REF. 2).
 CONFLICT 266 266 H -> H (IN REF. 4).
 CONFLICT 285 285 Y -> H (IN REF. 4).
 CONFLICT 301 302 DA -> EP (IN REF. 1).
 CONFLICT 310 310 L -> P (IN REF. 2).
 SEQUENCE 359 AA; 42123 MW; DD37176589E66046 CRC64;

 Query Match 53.1%; Score 1040; DB 1; Length 359;
 Best Local Similarity 55.8%; Pred. No. 2.1e-76;
 Matches 202; Conservative 65; Mismatches 83; Indels 12; Gaps 2;
 QY 13 CLTEDEKAARVDQENIRILLBQKQDRGELKLLLLGPGESGKSTFIKQMRIHAGYSE 72
 DB 10 CLSDEVKESKRINAETEKOLRRDKRDARRELKLLLLGTGSGKSTFIKQMRIHAGYSE 69
 QY 73 EERKGFRLPYQNIYFVMSRAMTEAMERLQIPSRPESKHHSLVMSODPKYVTTPEKRYA 132
 DB 70 EDKRGFTKLVIQNIITAMQAMIRAMETLKILYKYEONKANALLIREVDVKEVITFEHQIV 129
 QY 133 AAMQNLWRDAGTRACYERREPHLLDSAVYXLSHLERITEEGYVPTAQDVLRSRMTTGI 192
 DB 130 SAKLTWEDPGIQECYDRREYQLSDSAKYILTDVDRATLGLTLPQDVLKRVPTTGI 189
 QY 193 NEYCSVQKTNLRIVYDVGQSKERKWTCHFNVTALYLSLSEYDQCLENNQENRMK 252
 DB 190 IEYFPDLENIFRWYDVGQSKERRKWTCHFNVTALYLSLSEYDQCLENNQENRM 249
 QY 253 ESLALFGTILELPWPKSVSLFLNKTIDLEKIPTSHTLATYFPFGQPKQDAEAARFI 312
 DB 250 ESKALFRITIIYFPQNSVSLFLNKTIDLEKILYSLVDFPFDFGQPDQAAREFI 309
 QY 313 LDMYTRMTYTCVGDGEGSKGARSRLFSHYTCATDTONIRKRVKDVDRDVSRLARYLDEIN 372

Db 310 LKMEVDL-----NPDSK-----IIYSHFTCATDTENIRFVFAAVKDTIOLNLIKEYN 357
 QY 373 LL 374
 Db 358 LV 359

 RESULT 11
 ID CBLL_MOUSE STANDARD; PRT; 359 AA.
 AC P21278; Q61939;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 GN GUANINE NUCLEOTIDE-BINDING PROTEIN, 'ALPHA-11' SUBUNIT.
 GN GNALL OR GNA-11.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91067657; PubMed=2123549;
 RA Strathmann M., Simon M.I.;
 RT "G protein diversity: a distinct class of alpha subunits is present
 RT in vertebrates and invertebrates.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:9111-9117(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=129/SV;
 RC MEDLINE=96435432; PubMed=8838318;
 RA Davignon I., Barnard M., Gavrilova O., Sweet K., Wilkie T.M.;
 RT "Gene structure of murine Gna11 and Gna15: tandemly duplicated Gq
 RT class G protein alpha subunit genes.";
 RL Genomics 31:359-366(1996).
 RN [3]
 RP SEQUENCE OF 211-271 FROM N.A.
 RX MEDLINE=90017488; PubMed=2508088;
 RA Strathmann M., Wilkie T.M., Simon M.I.;
 RT "Diversity of the G-protein family: sequences from five additional
 RT alpha subunits in the mouse.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:7407-7409(1989).
 CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
 INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
 SIGNALING SYSTEMS. ACTS AS AN ACTIVATOR OF PHOSPHOLIPASE C.
 CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
 CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
 CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch)

 EMBL: M55411; AAA63305.1;
 EMBL: U37413; AAB36839.1;
 EMBL: U37411; AAB36839.1; JOINED.
 EMBL: U37412; AAB36839.1; JOINED.
 EMBL: M57617; AAA63301.1;
 PIR: B38414; RGM511.
 PIR: B38333; B38333.
 HSSP: P04896; LAZT.
 MGD: MGI:95766; Gna11.
 InterPro: IPR000654;
 InterPro: IPR001019;
 Pfam: PF00503; G-alpha; 1.
 PRINTS: PR00318; GPROTEINA.
 PRINTS: PR00442; GPROTEINAQ.
 GTP-binding; Transducer; Multigene family; ADP-ribosylation.
 KW

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; U89139; AAB49314.1; -
CC InterPro; IPR000654; -
CC InterPro; IPR001019; -
CC Pfam; PF00503; G-alpha; 1.
CC PRINTS; PR00318; GPROTEINA.
CC DR PRINTS; PR00442; GPROTEINAQ.
CC KW GTP-binding; Transducer; Multigene family; ADP-ribosylation;
KW Palmitate; Lipoprotein.
FT LIPID 3 3 PALMITATE (BY SIMILARITY).
FT LIPID 4 4 GTP (BY SIMILARITY).
FT NP_BIND 40 47 GTP (BY SIMILARITY).
FT NP_BIND 199 203 GTP (BY SIMILARITY).
FT NP_BIND 268 271 GTP (BY SIMILARITY).
FT MOD_RES 177 177 ADP-RIBOSYL[1] (BY ACTION OF CTX)
FT (BY SIMILARITY).
SQ SEQUENCE 353 AA; 41521 MW; 218B92CA16540408 CRC64;

Query Match 52.9%; Score 1036; DB 1; Length 353;
Best Local Similarity 56.4%; Pred. No. 4.4e-76;
Matches 204; Conservative 59; Mismatches 87; Indels 12; Gaps 2;

QY 13 CLTEDEAAARVDQINRILLEQKQKQDRELKLLGPGESGKSTFIKQRIIHGAGYSE 72
DB 4 CLSBEAKQKRIKNOIEIQLKRDARRELKLLGPGESGKSTFIKQRIIHGAGYSD 63
QY 73 BERKGRPLVYQNIQFVSMRAMIEAMERLIQIPFSRPEKHHASLVMSQDPYKVTTFE 132
DB 64 EDKRGFKLVQNIQFVSMRAMIEAMERLIQIPFSRPEKHHASLVMSQDPYKVTTFE 123
QY 133 AAMWLWDAGIRACIYERREFHLLDSAVVYLHLERITEEGYVPTAQDVLRSRPTTGI 192
DB 124 TAMNSLWQDTGIQCHYDRREYQLTDSAKYLLTDLRIAADKYVSTLQDILVRPTTGI 183
QY 193 NEYCFVQKTNLRIVDVGQKSRKWKHCFENVTALYLSLEYDQCLLENNOENRMK 252
DB 184 IEYFDLEIRFMRVVDVGQSRERKWKHCFENVTALYLSLEYDQCLLENNOENRM 243
QY 253 ESALFGTILELPWFKSTSVILFNKTDILEKPTSHLATYFPSPFGPKQDAEAKRFI 312
DB 244 ESKALFKTIITYPWFQHSVILFNKTDILEKPTSHLATYFPSPFGPKQDAEAKRFI 303
QY 313 LDMYTMVTCVDPGEGSKGARSRLFSHYTCATDTQNIKRVKDVDRSVLARYLDEIN 372
DB 304 LRMFVEL-----NPDPEK-----IIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYN 351
QY 373 LL 374
DB 352 LV 353

RESULT 13
GB11_MELGA STANDARD; PRT; 359 AA.
ID GB11_MELGA
AC P45645;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE GUANINE NUCLEOTIDE-BINDING PROTEIN, ALPHA-11 SUBUNIT.
GN GNAL1.
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Meleagridae; Meleagris.
OX NCBI_TaxID=9103;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=93207527; PubMed=8457205;

Thu Sep 6 15:16:58 2001

FT NP_BIND 46 53 GTP (BY SIMILARITY).
FT NP_BIND 205 209 GTP (BY SIMILARITY).
FT NP_BIND 274 277 GTP (BY SIMILARITY).
FT MOD_RES 183 183 ADP-RIBOSYL[1] (BY ACTION OF CTX)
FT (BY SIMILARITY).
FT CONFLICT 77 77 K -> L (IN REF. 2).
SQ SEQUENCE 359 AA; 42024 MW; A33D2D6C6C62F8D5 CRC64;

Query Match 53.0%; Score 1038; DB 1; Length 359;
Best Local Similarity 55.5%; Pred. No. 3.1e-76;
Matches 201; Conservative 65; Mismatches 84; Indels 12; Gaps 2;

QY 13 CLTEDEAAARVDQINRILLEQKQKQDRELKLLGPGESGKSTFIKQRIIHGAGYSE 72
DB 10 CLSDEVKESKRINAEIEQLKRDARRELKLLGPGESGKSTFIKQRIIHGAGYSE 69
QY 73 BERKGRPLVYQNIQFVSMRAMIEAMERLIQIPFSRPEKHHASLVMSQDPYKVTTFE 132
DB 70 EDKRGFKLVYQNIQFVSMRAMIEAMERLIQIPFSRPEKHHASLVMSQDPYKVTTFE 129
QY 133 AAMWLWDAGIRACIYERREFHLLDSAVVYLHLERITEEGYVPTAQDVLRSRPTTGI 192
DB 130 NAIKLWSDPGVQECYDRRERFQLSDAKYLLTDVRIATVGYLPTQDVLRVPTTGI 189
QY 193 NEYCFVQKTNLRIVDVGQKSRKWKHCFENVTALYLSLEYDQCLLENNOENRMK 252
DB 190 IEYFDLEIRFMRVVDVGQSRERKWKHCFENVTALYLSLEYDQCLLENNOENRM 249
QY 253 ESALFGTILELPWFKSTSVILFNKTDILEKPTSHLATYFPSPFGPKQDAEAKRFI 312
DB 250 ESKALFKTIITYPWFQHSVILFNKTDILEKPTSHLATYFPSPFGPKQDAEAKRFI 309
QY 313 LDMYTMVTCVDPGEGSKGARSRLFSHYTCATDTQNIKRVKDVDRSVLARYLDEIN 372
DB 310 LKMFVDEL-----NPDSDK-----IIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYN 357
QY 373 LL 374
DB 358 LV 359

RESULT 12
GBQ_HOMAM STANDARD; PRT; 353 AA.
ID GBQ_HOMAM
AC P91950;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE GUANINE NUCLEOTIDE-BINDING PROTEIN G(Q), ALPHA SUBUNIT.
OS Homarus americanus (American lobster).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
OC Nephropidae; Nephropidae; Homarus.
OX NCBI_TaxID=6706;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Olfactory organ;
RX MEDLINE=97309301; PubMed=9166716;
RA McClintock T.S., Xu F., Quintero J., Gress A.M., Landers T.M.;
RT "Molecular cloning of a lobster G alpha(q) protein expressed in
neurons of olfactory organ and brain."
RL J. Neurochem. 68:2248-2254(1997).
CC -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
SIGNALING SYSTEMS.
CC -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

QY 121 PYKVTTFEKRYAAAMQWLRRDAGIRACYERRERFHLDSAVYLSHLERITEEGVPTAQ 180
DB 118 VEKVCCTFEQYVNAIKNLWSDPGIOECYDRREYQLSDSTKYLLTDVDRISKPGYLPQQ 177
QY 181 DVLRSRMPPTTGINEYCFVQKTNLRIVDVGOKSKRWKHCFFENVIALIYLASEYDQ 240
DB 178 DVLVRVPPTGIIEXFPDLENIIFRMVDVGGORSEKRWKHCFFENVTSIMFLVALSEYDQ 237
QY 241 CLEENNOENRKESLALFCTILELPWFKSTSVILFUNKTDILEEKIPTSHLATYFPFSG 300
DB 238 VLVESDNENRMEESKALFRTIITYPWFQNSVILFUNKDLLEDKIMYSHLDVYPPFDG 297
QY 301 PKODAAAKRFILDMYTRMYTCGVDPGEGSKGARSRRFLFSHYTCATDQNRKVKFDYR 360
DB 298 PQDAATAREFILKMFVDL-----NPDSK-----IYSHFTCATDTEIRFVFAAVK 345
QY 361 DSVIARYLDEINLL 374
DB 346 DTILQHNKKEYNLV 359

Search completed: September 6, 2001, 10:55:23
Job time: 103 sec

Db 358 LV 359
ID GB11_XENLA STANDARD; PRT: 359 AA.
AC P43444;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE GUANINE NUCLEOTIDE-BINDING PROTEIN, ALPHA-II SUBUNIT.
GN Xa11.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocyte;
RX MEDLINE=94298961; PubMed=8026589;
RA Shapira H., Way J., Lipinsky D., Oron Y., Battey J.F.;
RT "Neuromedin B receptor, expressed in xenopus laevis oocytes,
selectively couples to G alpha q and not G alpha 11.";
RL FEBS Lett. 348:89-92(1994).
RN [2]
RP ERRATUM
RA Shapira H., Way J., Lipinsky D., Oron Y., Battey J.F.;
RL FEBS Lett. 349:318-318(1994).
CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
SIGNALING SYSTEMS. ACTS AS AN ACTIVATOR OF PHOSPHOLIPASE C.
CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(Q)).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U10494; AAA52188.1; -
CC HSSP; P04896; 1A2T.
DR InterPro: IPR000654; -
DR InterPro: IPR001019; -
DR Pfam: PF00503; G-alpha: 1.
DR PRINTS; PR00318; GPROTEINA.
DR PRINTS; PR00442; GPROTEINAQ.
KW GTP-binding; Transducer; Multigene family; ADP-ribosylation.
FT NP_BIND 46 53 GTP (BY SIMILARITY).
FT NP_BIND 205 209 GTP (BY SIMILARITY).
FT NP_BIND 274 277 GTP (BY SIMILARITY).
FT MOD_RES 183 183 ADP-RIBOSYL[1] (BY ACTION OF CTX) (BY
SIMILARITY).
FT SEQUENCE 359 AA; 42088 MW; DA6B376993FDD870 CRC64;
FT SQ

Query Match 52.4%; Score 1026.5; DB 1; Length 359;
Best Local Similarity 54.0%; Pred. No. 2.6e-75;
Matches 202; Conservative 65; Mismatches 92; Indels 15; Gaps 3;

QY 1 MARSLTWRCCPWCWTEDEKAARVDQENRILLQKQDRCELKLLLLGPGESGKSTFIK 60
DB 1 MTLDSTMAC---CLSEEVKSKRNAEIEKQLRRDKDRRELKLLLLGTGSGKSTFIK 57

QY 61 QMRITHGAGYSFEERKGFRLVYQNFVMSRAMTEAMERLQIPFSRPSKHHASLVMSQD 120
DB 58 QMRITHGSGSEEDKKGFTKLQVNFQIFTAMQSMIRAMETLKLILYKYEONKANAQVREVD 117

Thu Sep 6 15:16:58 2001

us-08-878-801-2.rsp

Page 12


```

Db 121 PYKNTPEYALAVQSLWLDAGVACRYERRERHLLDSAVYLSHLRIAEYGYPTAQ 180
QY 181 DVLSRMPPTGINEXCFVQKTNLRVDVGGQKSKRWKWHCFENVIALYLSLSEYDQ 240
Db 181 DVLSRMPPTGINEXCFVQKTNLRVDVGGQKSKRWKWHCFENVIALYLSLSEYDQ 240
QY 241 CLEENQENRMKESALFGLTLELFWFKSTSVILFUNKTDILEEKIPTSHLATYFPFSG 300
Db 241 CLEENQENRMKESALFGLTLELFWFKSTSVILFUNKTDILEEKIPTSHLATYFPFSG 300
QY 301 PQDAAARFILDYRMVTCVDPGEGSKGARSRLFSHYTCATDTQNIKRVKVDVR 360
Db 301 PQDAAARFILDYRMVTCVDPGEGSKGARSRLFSHYTCATDTQNIKRVKVDVR 360
QY 361 DSVLYRLDEINLL 374
Db 361 DSVLYRLDEINLL 374

RESULT 2
ID Q9U473 PRELIMINARY; PRT; 353 AA.
AC Q9U473;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE G-ALPHA-49B OR CG17759.
OS Panulirus argus (Spiny lobster).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Palinura;
OC Palinuroidea; Palinuridae; Panuliridae; Panulirus.
NCBI_TaxID=6737;
[1]
SEQUENCE FROM N.A.
RC TSSUB-OLFACTORY ORGAN;
RA Mungar S.D., Gleeson R.A., Aldrich H.C., Rust N.C., Ache B.W.,
RA Greenberg R.M.;
RT "Molecular Evidence for Phosphoinositide-Mediated Signaling in Lobster
RT Olfactory Receptor Neurons."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF201328; AAF19378.1;
DR HSP: P10824; 180F.
DR InterPro: IP001019;
DR InterPro: IP001019;
DR Pfam: PF00503; G-alpha; 1.
DR PRINTS: PR00318; GPROTEINA.
DR ProDom: PD001018; 1.
DR SMART: SM00275; G-alpha; 1.
SQ
SEQUENCE 353 AA; 41478 MW; FD9B551F66327BAA CRC64;

Query Match 52.6%; Score 1031; DB 5; Length 353;
Best Local Similarity 56.1%; Pred. No. 1.8e-77;
Matches 203; Conservative 59; Mismatches 88; Indels 12; Gaps 2;

QY 13 CLTEDEKAAARVDQETNRIILPOKKDRGELKLLGLPGESGKSTFIKMRHICAGYSE 72
Db 4 CLSEAKEQKRIHQETNRIILPOKKDRGELKLLGLPGESGKSTFIKMRHICAGYSD 63
QY 73 EERKGFRLVYONIFVSRMATEAMERLQIPSPESKHASLVMSQDPYKVTTFERYA 132
Db 64 EDRKGFRLVYONIFVSRMATEAMERLQIPSPESKHASLVMSQDPYKVTTFERYA 123
QY 133 AMQWLWDAGTRACYRRERHLLDSAVYLSHLRIAEYGYPTAQDVLRSRMPPTGI 192
Db 124 TAMKSLWQDTQICQYDRREYQITDSAKYLLTDLDRIAATDYVSTLQDILRVRAPTGI 183
QY 193 NEYCFVQKTNLRVDVGGQKSKRWKWHCFENVIALYLSLSEYDQCLEENQENRMK 252
Db 184 IXPFDLEEIRFMDVGGQKSKRWKWHCFENVIALYLSLSEYDQILEENQENRMK 243
QY 253 ESLALFTGILPFWFKSTSVILFUNKTDILEEKIPTSHLATYFPFSGQPKQDAEAAKRFI 312

```

```

Db 244 ESKALFKTIIITYPFQHSVILFNKKDLLEKIMYSHLVDFPEYDQPRKDAIAAREFI 303
QY 313 LDMYTRMYTCVDPGEGSKGARSRLFSHYTCATDTQNIKRVKVDVRSVLYRLDEIN 372
Db 304 LRMFVEL-----NPDPEK-----IYSHF-CATDTENIRFVFAAVKDTILQLNLKEYN 351
QY 373 LL 374
Db 352 LV 353

RESULT 3
ID Q9V6E2 PRELIMINARY; PRT; 353 AA.
AC Q9V6E2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE G-ALPHA-49B OR CG17759.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
[1]
SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.G., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.T., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Bozova D., Botchan M.R., Bouck J., Borktein P., Bottler P.,
RA Butts J.M., Busam D.A., Butler H., Caidle E., Center A., Chandra I.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Ke Z., Kennison J.A., Ketchum K.A.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy J., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stappleton M., Strong R., Sun E.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.E., Wang A.H., Wang X.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu S., Zhu S.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Verter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AF003821; AAF58485.2;
DR HSP: P10824; 180F.
DR FlyBase: FBgn0004435; G-alpha-49B.
DR InterPro: IPR001019;
DR Pfam: PF00503; G-alpha; 2.

```

1

Db 184 LEYFPDLGIVFMDVGGQSRERRKWHICFENVNTSIIFLVALSEYDQILFSDNENRME 243
 QY 253 ESALFGTILELPWKSTSVILFNKTDILEKIPSTSHLATYFPFGQKODAAAKRFI 312
 Db 244 ESALFRTIITPWFONSSVILFNKTDILEKIMSHLVDFPEYDGPCKDHAARQFV 303
 QY 313 LDMYRMTYTCVGGEGSKGARSRLSHYTCATDTONIRKVKDVRSDSVLARYLDEIN 372
 Db 304 L-----KKYLAANPDPE-----RQCYSHETTATDTENIKLVCAVKDTIMQNALKEFN 351
 QY 373 L 373
 Db 352 L 352
 RESULT 6
 ID P91955 PRELIMINARY: PRT: 353 AA.
 AC P91955;
 DT 01-MAY-1997 (Tremblrel. 03, Created)
 DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE GQ PROTEIN ALPHA SUBUNIT
 OS Limulus polyphemus (Atlantic horseshoe crab)
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;
 OC Limulidae; Limulus.
 OX NCBI_TaxID=6850;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=VENTRAL EYE;
 RA Mungier S.D., Schrenser-Berlin J.L., Brink C.M., Battelle B.A.;
 RA Invert. Neurosci. 0:0-0(0).
 DR EMBL; U88586; AAB48510.1; -;
 DR HSP; P10824; IBOF.
 DR InterPro; IPR001019; -;
 DR Pfam; PF00503; G-alpha; 1.
 DR PRINTS; PR00318; GPROTEINA.
 DR SMART; SM00275; G-alpha; 1.
 SQ SEQUENCE 353 AA; 41497 MW; C1C57783B3D2D516 CRC64;

Query Match 50.3%; Score 986; DB 5; Length 353;
 Best Local Similarity 53.9%; Pred. No. 9.5e-74;
 Matches 195; Conservative 62; Mismatches 93; Indels 12; Gaps 2;
 QY 13 CLTEDEKAAARVDQENRILLEOKKODRGELKLLILGPGESGKSTFIKQRIIHGAGYSE 72
 Db 4 CLSEEGKEQKRINQETEROLRDKRDARRELKLLILGTGSGCKSTFIKQRIIHGAGYSD 63
 QY 73 EERKGRPLVYQNIYFMSRAMIEMERLQIPFSRPSKHHASLVMSQDPYKVTTFEKR 132
 Db 64 DKKSYKILVYQNIYFMSRAMIEMERLQIPFSRPSKHHASLVMSQDPYKVTTFEKR 132
 QY 133 AMQWLWRDAGIRACVYRRREFHLLDSAVYILSHLERITEEGYVPTAODVLRSMPTTGI 192
 Db 124 EAKLSLVDPGEGEYDRREYQLTDSAKYLLINDIDRIVPNYLPQDILRVVRPTTGI 183
 QY 193 NPYCSVQKTNIRIVDVGGQSKERRKWHICFENVNTSIIFLVALSEYDQILFSDNENRME 252
 Db 184 IEYFILDIIIFMDVGGQSRERRKWHICFENVNTSIIFLVALSEYDQILFSDNENRME 243
 QY 253 ESALFGTILELPWKSTSVILFNKTDILEKIPSTSHLATYFPFGQKODAAAKRFI 312
 Db 244 ESALFRTIITPWFONSSVILFNKTDILEKIMSHLVDFPEYDGPCKDHAARQFV 303
 QY 313 LDMYRMTYTCVGGEGSKGARSRLSHYTCATDTONIRKVKDVRSDSVLARYLDEIN 372
 Db 304 LMFVDL-----NPDSEK-----IYSHFTCATDTENIRFVFAAVKDTILQLNLKEYN 351
 QY 373 LL 374
 Db 352 LV 353

RESULT 7
 ID Q17386 PRELIMINARY: PRT: 355 AA.
 AC Q17386; 002546;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE EGL-30.
 GN EGL-30 OR W01D7.7
 OS Caenorhabditis elegans
 OC Eukaryota; Metazoa; Hematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N2;
 RC MEDLINE=96221161; PubMed=8630258;
 RA Brundage L., Avery L., Katz A., Kim U.J., Mendel J.E., Sternberg P.W.,
 RA Simon M.I.;
 RA "Mutations in a C. elegans Galpha gene disrupt movement, egg laying,
 RA and viability";
 RA Neuron 16:999-1009(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RC MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berke M., Coulson A.,
 RA Bonfield J., Burton J., Connell M., Gopsey T., Cooper J., Favello A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roop A., Saunders D., Shownkeen R.,
 RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J., Waterston R.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,
 RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RA elegans";
 RA Nature 368:32-38(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Gattung S., Goela D., Wilson R.;
 RA Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RA Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U56864; AAB04059.1; -;
 DR EMBL; AF003739; AAB58071.1; -;
 DR HSP; P10824; IBOF.
 DR InterPro; IPR001019; -;
 DR Pfam; PF00503; G-alpha; 1.
 DR PRINTS; PR00318; GPROTEINA.
 DR SMART; SM00275; G-alpha; 1.
 SQ SEQUENCE 355 AA; 41865 MW; FE438B01C2E1355C CRC64;
 Query Match 49.7%; Score 974; DB 5; Length 355;
 Best Local Similarity 53.0%; Pred. No. 9.4e-73;
 Matches 193; Conservative 64; Mismatches 93; Indels 14; Gaps 3;
 QY 13 CLTEDEKAAARVDQENRILLEOKKODRGELKLLILGPGESGKSTFIKQRIIHGAGYSE 72
 Db 4 CLSEEGKEQKRINQETEROLRDKRDARRELKLLILGTGSGCKSTFIKQRIIHGAGYSE 63
 QY 73 EERKGRPLVYQNIYFMSRAMIEMERLQIPFSRPSKHHASLVMSQDPYKVTTFEKR 130
 Db 64 EDRKGRPLVYQNIYFMSRAMIEMERLQIPFSRPSKHHASLVMSQDPYKVTTFEKR 123
 QY 131 YAAAMQWLWRDAGIRACVYRRREFHLLDSAVYILSHLERITEEGYVPTAODVLRSMPTT 190

```

Db 124 VYSYIKELWEDSGQECVDRREYQLTDSAXYLSDLRLAVDPYLTPEQDILRVPTT 183
Qy 191 GINCYFVQKTNLRIVDVGQKSRKWIHCFFENFVIALYSLSEYDQCLENNQENR 250
Db 184 GIIEYFDFLEQIIPFVMDVGGORSEKRWIHCFFENVTSMELVALSEYDQVLECDNENR 243
Qy 251 MKESALFSGTILELPWFKSTSVILFLNKTDLLEKIPSHLATYFPPSGQPKQDAEAAR 310
Db 244 MEESKALFRTITYPWFNTSSVILFLNKKDLLEKILYSLADYFPEYDGGPRDPIARE 303
Qy 311 FILDYTRMYTCVDPGSGKRGARSLRFSHYTCATDTQNIKRVKFDVRSVLARYLDE 370
Db 304 FLKMFVDL-----NPDADK-----IYSHFTCATDTENIRFVFAAVKOTILQHNLKE 351
Qy 371 INLL 374
Db 352 YNLV 355

RESULT 8
Q9X2V4 PRELIMINARY; PRT; 355 AA.
AC Q9X2V4;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE GQ PROTEIN, ALPHA SUBUNIT.
GN GQ.
OS Geodia cydonium (Sponge).
OC Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha;
OC Astrophorida; Geodiidae; Geodia.
OX NCBI_TaxID=6047;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98119577; PubMed=9459489;
RA Seack J., Kruse M., Mueller W.E.G.;
RT "Evolutionary analysis of G-proteins in early metazoans: cloning of
alpha- and beta-subunits from the sponge Geodia cydonium.";
RL Biochim. Biophys. Acta 1401:93-103(1998).
DR EMBL; Y14248; CAB43527.1; -.
DR HSSP; P04896; LAZT.
DR InterPro; IPR001019; -.
DR Pfam; PF00503; G-alpha; 1.
DR PRINTS; PR00318; GPROTEINA.
DR SMART; SM00275; G-alpha; 1.
SQ SEQUENCE 355 AA; 41363 MW; 6828C29643F2CE91 CRC64;

```

```

Query Match 44.7%; Score 876.5; DB 5; Length 355;
Best Local Similarity 47.1%; Pred. No. 1.1e-64;
Matches 169; Conservative 73; Mismatches 104; Indels 13; Gaps 2;

Qy 14 LTEDKAAARVDQFINRILEQKQDQGEUKLLLLGSGSGKSTFTKQRIHGCYSEE 73
Db 5 LSEERLQKRINRINRELQDQKAKETKLLLLGTGSGSGKSTFTKQRIHGCYSEE 64
Qy 74 ERKGFRLVYQNFVSNRMIAEMERLQIPESRPSKHHASLVMSODPYKVTTFEKRYAA 133
Db 65 DCLEYKNLVFNILMSHMSLQTAELKAIYIDPDQRHVQLLMALRPETAQSLGGTCE 124
-Qy 134 AMQWLWDAGIRACRYERRRFFHLDSAVVYLSHLERITEGYPTAODVLRSMPTTGIN 193
Db 125 AIRKLWDAGVQCYQRNREYQLSDSTPKYLLDLPRISSNDYVPTTQDVLVRVPTTGIN 184
Qy 194 EYCSVQKTNLRIVDVGQKSRKWIHCFFENFVIALYSLSEYDQCLENNQENRMK 252
Db 185 EYPTINKIIFKMWVGGQSRSEKRWIHCFFDHVTSNMFVAISEYDQILVEADSVNRMV 244
Qy 253 ESLALFGTILELPWFKSTSVILFLNKTDLLEKIPSHLATYFPPSGQPKQDAEAARFI 312
Db 245 ESLHLFNTIISYPWFNKSIFLFLNKKDLLEKIVHSHLDIYFEEYDGPCKDHVSARESI 304

```

```

Qy 313 LDMYTRMYTCVDPGSGKRGARSLRFSHYTCATDTQNIKRVKFDVRSVLARYLDEI 371
Db 305 AKMFI-----SINDMRSADIYPHFTCATDTENIKFVDFVYVKNHILQOQHTEV 351

RESULT 9
Q9Y207 PRELIMINARY; PRT; 305 AA.
AC Q9Y207;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE G PROTEIN A SUBUNIT 3 (FRAGMENT).
OS Hydra magnipapillata (Hydra).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroids; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6085;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99246375; PubMed=10229568;
RA Suga H., Koyanagi M., Hoshiyama D., Ono K., Iwabe N., Kuma K.,
RA Miyata T.;
RT "Extensive gene duplication in the early evolution of animals before
the parazoan-eumetazoan split demonstrated by G proteins and protein
tyrosine kinases from sponge and hydra.";
RL J. Mol. Evol. 48:646-653(1999).
DR EMBL; AB006541; BAA81695.1; -.
DR HSSP; P10824; IBOF.
DR InterPro; IPR001019; -.
DR Pfam; PF00503; G-alpha; 1.
DR PRINTS; PR00318; GPROTEINA.
DR SMART; SM00275; G-alpha; 1.
FT NON_TER 1
SQ SEQUENCE 305 AA; 35491 MW; 2639884D438586E9 CRC64;

```

```

Query Match 42.3%; Score 829.5; DB 5; Length 305;
Best Local Similarity 52.1%; Pred. No. 7.1e-61;
Matches 165; Conservative 57; Mismatches 78; Indels 17; Gaps 4;

Qy 63 RIHHCAGYSEERKGFRLVYQNFVSNRMIAEMERLQIPESRPSKHHASLV-----MS 118
Db 1 RIHHCQGVTEDDRGTNLVFLNLIYQALTRAMNLKISYSPANEENARLLDVLDS 60
Qy 119 QDPYKVT-TFEKRYAAAMQWLWDAGIRACRYERRRFFHLDSAVVYLSHLERITEGYVP 177
Db 61 EDKSRITTLTSPQASAIESLWKDSCGIEVDYDRREYQLSDSAKYLSDLKRICAPNVP 120
Qy 178 TAQDVLRSRMPPTGINEYCFVQKTNLRIVDVGQKSRKWIHCFFENFVIALYSLSE 237
Db 121 TMQDVLRAAPTGTGIIIEYFEDLDTIIFRMVDVGGQSRSEKRWIHCFFENVTSMFLVALSE 180
Qy 238 YDQCLENNQENRMKESALFSGTILELPWFKSTSVILFLNKTDLLEKIPSHLATYTPPS 297
Db 181 YDQVLFESQENRMDSEKALFTIITYPWFQSSIIILFNKTDLLEKIQKSDLTQYFPE 240
Qy 298 FQGPQDAAEAKRFILDMYTRMYTCVDPGSGKRGARSLRFSHYTCATDTQNIKRVFK 357
Db 241 YDQAGKDAKAAEFILKMFVDL-----NPDYDK-----IYSHFTCATDTENIRFVFA 288
Qy 358 DYRDSVLARYLDEINLL 374
Db 289 AVRDITLQNLKYNLV 305

RESULT 10
Q9I7C8 PRELIMINARY; PRT; 279 AA.
AC Q9I7C8;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE G PROTEIN ALPHA 49B.

```

GN G-ALPHA-49B OR CG17759.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
-RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhattacharya S., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Rosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaiswal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Fairclat R., Remington K., Saunders R.D.C., Scheeler F., Smith T.,
RA Shue B.C., Siding-Klamis I., Simpson M., Skupski M.P., Shen T.,
RA Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Strickas R., Turner R., Ventner E., Wang A.H., Wang X.,
RA Wang J., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Will J., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Zhang R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zeng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "the genome of Drosophila melanogaster";
RT Science 287:2185-2195 (2000).
RL EMBL: AF003821, AF003822, AF003823, AF003824, AF003825, AF003826, AF003827, AF003828, AF003829, AF003830, AF003831, AF003832, AF003833, AF003834, AF003835, AF003836, AF003837, AF003838, AF003839, AF003840, AF003841, AF003842, AF003843, AF003844, AF003845, AF003846, AF003847, AF003848, AF003849, AF003850, AF003851, AF003852, AF003853, AF003854, AF003855, AF003856, AF003857, AF003858, AF003859, AF003860, AF003861, AF003862, AF003863, AF003864, AF003865, AF003866, AF003867, AF003868, AF003869, AF003870, AF003871, AF003872, AF003873, AF003874, AF003875, AF003876, AF003877, AF003878, AF003879, AF003880, AF003881, AF003882, AF003883, AF003884, AF003885, AF003886, AF003887, AF003888, AF003889, AF003890, AF003891, AF003892, AF003893, AF003894, AF003895, AF003896, AF003897, AF003898, AF003899, AF003900, AF003901, AF003902, AF003903, AF003904, AF003905, AF003906, AF003907, AF003908, AF003909, AF003910, AF003911, AF003912, AF003913, AF003914, AF003915, AF003916, AF003917, AF003918, AF003919, AF003920, AF003921, AF003922, AF003923, AF003924, AF003925, AF003926, AF003927, AF003928, AF003929, AF003930, AF003931, AF003932, AF003933, AF003934, AF003935, AF003936, AF003937, AF003938, AF003939, AF003940, AF003941, AF003942, AF003943, AF003944, AF003945, AF003946, AF003947, AF003948, AF003949, AF003950, AF003951, AF003952, AF003953, AF003954, AF003955, AF003956, AF003957, AF003958, AF003959, AF003960, AF003961, AF003962, AF003963, AF003964, AF003965, AF003966, AF003967, AF003968, AF003969, AF003970, AF003971, AF003972, AF003973, AF003974, AF003975, AF003976, AF003977, AF003978, AF003979, AF003980, AF003981, AF003982, AF003983, AF003984, AF003985, AF003986, AF003987, AF003988, AF003989, AF003990, AF003991, AF003992, AF003993, AF003994, AF003995, AF003996, AF003997, AF003998, AF003999, AF004000, AF004001, AF004002, AF004003, AF004004, AF004005, AF004006, AF004007, AF004008, AF004009, AF004010, AF004011, AF004012, AF004013, AF004014, AF004015, AF004016, AF004017, AF004018, AF004019, AF004020, AF004021, AF004022, AF004023, AF004024, AF004025, AF004026, AF004027, AF004028, AF004029, AF004030, AF004031, AF004032, AF004033, AF004034, AF004035, AF004036, AF004037, AF004038, AF004039, AF004040, AF004041, AF004042, AF004043, AF004044, AF004045, AF004046, AF004047, AF004048, AF004049, AF004050, AF004051, AF004052, AF004053, AF004054, AF004055, AF004056, AF004057, AF004058, AF004059, AF004060, AF004061, AF004062, AF004063, AF004064, AF004065, AF004066, AF004067, AF004068, AF004069, AF004070, AF004071, AF004072, AF004073, AF004074, AF004075, AF004076, AF004077, AF004078, AF004079, AF004080, AF004081, AF004082, AF004083, AF004084, AF004085, AF004086, AF004087, AF004088, AF004089, AF004090, AF004091, AF004092, AF004093, AF004094, AF004095, AF004096, AF004097, AF004098, AF004099, AF004100, AF004101, AF004102, AF004103, AF004104, AF004105, AF004106, AF004107, AF004108, AF004109, AF004110, AF004111, AF004112, AF004113, AF004114, AF004115, AF004116, AF004117, AF004118, AF004119, AF004120, AF004121, AF004122, AF004123, AF004124, AF004125, AF004126, AF004127, AF004128, AF004129, AF004130, AF004131, AF004132, AF004133, AF004134, AF004135, AF004136, AF004137, AF004138, AF004139, AF004140, AF004141, AF004142, AF004143, AF004144, AF004145, AF004146, AF004147, AF004148, AF004149, AF004150, AF004151, AF004152, AF004153, AF004154, AF004155, AF004156, AF004157, AF004158, AF004159, AF004160, AF004161, AF004162, AF004163, AF004164, AF004165, AF004166, AF004167, AF004168, AF004169, AF004170, AF004171, AF004172, AF004173, AF004174, AF004175, AF004176, AF004177, AF004178, AF004179, AF004180, AF004181, AF004182, AF004183, AF004184, AF004185, AF004186, AF004187, AF004188, AF004189, AF004190, AF004191, AF004192, AF004193, AF004194, AF004195, AF004196, AF004197, AF004198, AF004199, AF004200, AF004201, AF004202, AF004203, AF004204, AF004205, AF004206, AF004207, AF004208, AF004209, AF004210, AF004211, AF004212, AF004213, AF004214, AF004215, AF004216, AF004217, AF004218, AF004219, AF004220, AF004221, AF004222, AF004223, AF004224, AF004225, AF004226, AF004227, AF004228, AF004229, AF004230, AF004231, AF004232, AF004233, AF004234, AF004235, AF004236, AF004237, AF004238, AF004239, AF004240, AF004241, AF004242, AF004243, AF004244, AF004245, AF004246, AF004247, AF004248, AF004249, AF004250, AF004251, AF004252, AF004253, AF004254, AF004255, AF004256, AF004257, AF004258, AF004259, AF004260, AF004261, AF004262, AF004263, AF004264, AF004265, AF004266, AF004267, AF004268, AF004269, AF004270, AF004271, AF004272, AF004273, AF004274, AF004275, AF004276, AF004277, AF004278, AF004279, AF004280, AF004281, AF004282, AF004283, AF004284, AF004285, AF004286, AF004287, AF004288, AF004289, AF004290, AF004291, AF004292, AF004293, AF004294, AF004295, AF004296, AF004297, AF004298, AF004299, AF004300, AF004301, AF004302, AF004303, AF004304, AF004305, AF004306, AF004307, AF004308, AF004309, AF004310, AF004311, AF004312, AF004313, AF004314, AF004315, AF004316, AF004317, AF004318, AF004319, AF004320, AF004321, AF004322, AF004323, AF004324, AF004325, AF004326, AF004327, AF004328, AF004329, AF004330, AF004331, AF004332, AF004333, AF004334, AF004335, AF004336, AF004337, AF004338, AF004339, AF004340, AF004341, AF004342, AF004343, AF004344, AF004345, AF004346, AF004347, AF004348, AF004349, AF004350, AF004351, AF004352, AF004353, AF004354, AF004355, AF004356, AF004357, AF004358, AF004359, AF004360, AF004361, AF004362, AF004363, AF004364, AF004365, AF004366, AF004367, AF004368, AF004369, AF004370, AF004371, AF004372, AF004373, AF004374, AF004375, AF004376, AF004377, AF004378, AF004379, AF004380, AF004381, AF004382, AF004383, AF004384, AF004385, AF004386, AF004387, AF004388, AF004389, AF004390, AF004391, AF004392, AF004393, AF004394, AF004395, AF004396, AF004397, AF004398, AF004399, AF004400, AF004401, AF004402, AF004403, AF004404, AF004405, AF004406, AF004407, AF004408, AF004409, AF004410, AF004411, AF004412, AF004413, AF004414, AF004415, AF004416, AF004417, AF004418, AF004419, AF004420, AF004421, AF004422, AF004423, AF004424, AF004425, AF004426, AF004427, AF004428, AF004429, AF004430, AF004431, AF004432, AF004433, AF004434, AF004435, AF004436, AF004437, AF004438, AF004439, AF004440, AF004441, AF004442, AF004443, AF004444, AF004445, AF004446, AF004447, AF004448, AF004449, AF004450, AF004451, AF004452, AF004453, AF004454, AF004455, AF004456, AF004457, AF004458, AF004459, AF004460, AF004461, AF004462, AF004463, AF004464, AF004465, AF004466, AF004467, AF004468, AF004469, AF004470, AF004471, AF004472, AF004473, AF004474, AF004475, AF004476, AF004477, AF004478, AF004479, AF004480, AF004481, AF004482, AF004483, AF004484, AF004485, AF004486, AF004487, AF004488, AF004489, AF004490, AF004491, AF004492, AF004493, AF004494, AF004495, AF004496, AF004497, AF004498, AF004499, AF004500, AF004501, AF004502, AF004503, AF004504, AF004505, AF004506, AF004507, AF004508, AF004509, AF004510, AF004511, AF004512, AF004513, AF004514, AF004515, AF004516, AF004517, AF004518, AF004519, AF004520, AF004521, AF004522, AF004523, AF004524, AF004525, AF004526, AF004527, AF004528, AF004529, AF004530, AF004531, AF004532, AF004533, AF004534, AF004535, AF004536, AF004537, AF004538, AF004539, AF004540, AF004541, AF004542, AF004543, AF004544, AF004545, AF004546, AF004547, AF004548, AF004549, AF004550, AF004551, AF004552, AF004553, AF004554, AF004555, AF004556, AF004557, AF004558, AF004559, AF004560, AF004561, AF004562, AF004563, AF004564, AF004565, AF004566, AF004567, AF004568, AF004569, AF004570, AF004571, AF004572, AF004573, AF004574, AF004575, AF004576, AF004577, AF004578, AF004579, AF004580, AF004581, AF004582, AF004583, AF004584, AF004585, AF004586, AF004587, AF004588, AF004589, AF004590, AF004591, AF004592, AF004593, AF004594, AF004595, AF004596, AF004597, AF004598, AF004599, AF004600, AF004601, AF004602, AF004603, AF004604, AF004605, AF004606, AF004607, AF004608, AF004609, AF004610, AF004611, AF004612, AF004613, AF004614, AF004615, AF004616, AF004617, AF004618, AF004619, AF004620, AF004621, AF004622, AF004623, AF004624, AF004625, AF004626, AF004627, AF004628, AF004629, AF004630, AF004631, AF004632, AF004633, AF004634, AF004635, AF004636, AF004637, AF004638, AF004639, AF004640, AF004641, AF004642, AF004643, AF004644, AF004645, AF004646, AF004647, AF004648, AF004649, AF004650, AF004651, AF004652, AF004653, AF004654, AF004655, AF004656, AF004657, AF004658, AF004659, AF004660, AF004661, AF004662, AF004663, AF004664, AF004665, AF004666, AF004667, AF004668, AF004669, AF004670, AF004671, AF004672, AF004673, AF004674, AF004675, AF004676, AF004677, AF004678, AF004679, AF004680, AF004681, AF004682, AF004683, AF004684, AF004685, AF004686, AF004687, AF004688, AF004689, AF004690, AF004691, AF004692, AF004693, AF004694, AF004695, AF004696, AF004697, AF004698, AF004699, AF004700, AF004701, AF004702, AF004703, AF004704, AF004705, AF004706, AF004707, AF004708, AF004709, AF004710, AF004711, AF004712, AF004713, AF004714, AF004715, AF004716, AF004717, AF004718, AF004719, AF004720, AF004721, AF004722, AF004723, AF004724, AF004725, AF004726, AF004727, AF004728, AF004729, AF004730, AF004731, AF004732, AF004733, AF004734, AF004735, AF004736, AF004737, AF004738, AF004739, AF004740, AF004741, AF004742, AF004743, AF004744, AF004745, AF004746, AF004747, AF004748, AF004749, AF004750, AF004751, AF004752, AF004753, AF004754, AF004755, AF004756, AF004757, AF004758, AF004759, AF004760, AF004761, AF004762, AF004763, AF004764, AF004765, AF004766, AF004767, AF004768, AF004769, AF004770, AF004771, AF004772, AF004773, AF004774, AF004775, AF004776, AF004777, AF004778, AF004779, AF004780, AF004781, AF004782, AF004783, AF004784, AF004785, AF004786, AF004787, AF004788, AF004789, AF004790, AF004791, AF004792, AF004793, AF004794, AF004795, AF004796, AF004797, AF004798, AF004799, AF004800, AF004801, AF004802, AF004803, AF004804, AF004805, AF004806, AF004807, AF004808, AF004809, AF004810, AF004811, AF004812, AF004813, AF004814, AF004815, AF004816, AF004817, AF004818, AF004819, AF004820, AF004821, AF004822, AF004823, AF004824, AF004825, AF004826, AF004827, AF004828, AF004829, AF004830, AF004831, AF004832, AF004833, AF004834, AF004835, AF004836, AF004837, AF004838, AF004839, AF004840, AF004841, AF004842, AF004843, AF004844, AF004845, AF004846, AF004847, AF004848, AF004849, AF004850, AF004851, AF004852, AF004853, AF004854, AF004855, AF004856, AF004857, AF004858, AF004859, AF004860, AF004861, AF004862, AF004863, AF004864, AF004865, AF004866, AF004867, AF004868, AF004869, AF004870, AF004871, AF004872, AF004873, AF004874, AF004875, AF004876, AF004877, AF004878, AF004879, AF004880, AF004881, AF004882, AF004883, AF004884, AF004885, AF004886, AF004887, AF004888, AF004889, AF004890, AF004891, AF004892, AF004893, AF004894, AF004895, AF004896, AF004897, AF004898, AF004899, AF004900, AF004901, AF004902, AF004903, AF004904, AF004905, AF004906, AF004907, AF004908, AF004909, AF004910, AF004911, AF004912, AF004913, AF004914, AF004915, AF004916, AF004917, AF004918, AF004919, AF004920, AF004921, AF004922, AF004923, AF004924, AF004925, AF004926, AF004927, AF004928, AF004929, AF004930, AF004931, AF004932, AF004933, AF004934, AF004935, AF004936, AF004937, AF004938, AF004939, AF004940, AF004941, AF004942, AF004943, AF004944, AF004945, AF004946, AF004947, AF004948, AF004949, AF004950, AF004951, AF004952, AF004953, AF004954, AF004955, AF004956, AF004957, AF004958, AF004959, AF004960, AF004961, AF004962, AF004963, AF004964, AF004965, AF004966, AF004967, AF004968, AF004969, AF004970, AF004971, AF004972, AF004973, AF004974, AF004975, AF004976, AF004977, AF004978, AF004979, AF004980, AF004981, AF004982, AF004983, AF004984, AF004985, AF004986, AF004987, AF004988, AF004989, AF004990, AF004991, AF004992, AF004993, AF004994, AF004995, AF004996, AF004997, AF004998, AF004999, AF005000, AF005001, AF005002, AF005003, AF005004, AF005005, AF005006, AF005007, AF005008, AF005009, AF005010, AF005011, AF005012, AF005013, AF005014, AF005015, AF005016, AF005017, AF005018, AF005019, AF005020, AF005021, AF005022, AF005023, AF005024, AF005025, AF005026, AF005027, AF005028, AF005029, AF005030, AF005031, AF005032, AF005033, AF005034, AF005035, AF005036, AF005037, AF005038, AF005039, AF005040, AF005041, AF005042, AF005043, AF005044, AF005045, AF005046, AF005047, AF005048, AF005049, AF005050, AF005051, AF005052, AF005053, AF005054, AF005055, AF005056, AF005057, AF005058, AF005059, AF005060, AF005061, AF005062, AF005063, AF005064, AF005065, AF005066, AF005067, AF005068, AF005069, AF005070, AF005071, AF005072, AF005073, AF005074, AF005075, AF005076, AF005077, AF005078, AF005079, AF005080, AF005081, AF005082, AF005083, AF005084, AF005085, AF005086, AF005087, AF005088, AF005089, AF005090, AF005091, AF005092, AF005093, AF005094, AF005095, AF005096, AF005097, AF005098, AF005099, AF005100, AF005101, AF005102, AF005103, AF005104, AF005105, AF005106, AF005107, AF005108, AF005109, AF005110, AF005111, AF005112, AF005113, AF005114, AF005115, AF005116, AF005117, AF005118, AF005119, AF005120, AF005121, AF005122, AF005123, AF005124, AF005125, AF005126, AF005127, AF005128, AF005129, AF005130, AF005131, AF005132, AF005133, AF005134, AF005135, AF005136, AF005137, AF005138, AF005139, AF005140, AF005141, AF005142, AF005143, AF005144, AF005145, AF005146, AF005147, AF005148, AF005149, AF005150, AF005151, AF005152, AF005153, AF005154, AF005155, AF005156, AF005157, AF005158, AF005159, AF005160, AF005161, AF005162, AF005163, AF005164, AF005165, AF005166, AF005167, AF005168, AF005169, AF005170, AF005171, AF005172, AF005173, AF005174, AF005175, AF005176, AF005177, AF005178, AF005179, AF005180, AF005181, AF005182, AF005183, AF005184, AF005185, AF005186, AF005187, AF005188, AF005189, AF005190, AF005191, AF005192, AF005193, AF005194, AF005195, AF005196, AF005197, AF005198, AF00

RC STRAIN-W8;

RA Aimi T., Sanae K., Wang Q., Morinaga T.;
 RT "Molecular cloning of three genes for G protein alpha-subunit-protein
 from white root rot fungus, *Rosellinia necatorix*.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB045578; BAB20820.1; -;
 SQ SEQUENCE 353 AA; 41016 MW; 0EB84C4B292B70E6 CRC64;

Query Match 39.0%; Score 764.5; DB 3; Length 353;
 Best Local Similarity 42.7%; Pred. No. 2.1e-55;
 Matches 152; Conservative 67; Mismatches 120; Indels 17; Gaps 5;

QY 15 TEDEKAARVDQENRILLQKQKDRGELKLLLLGPGESGKSTFIKQMRHHGAGYSEEE 74
 DB 7 TEEKEGAR-NEETENQKRDQKMMORNEIKMLLLGAGESGKSTILKQMKLHIEGGYSRDE 65
 QY 75 RKGFRLVYQNIQFVSMRAMIEMERLQIPFSRPSKHHASIVMSQDPYK--TTFEKRYA 132
 DB 66 RESFKEIFSNVQSMRVLAEESLEPLPEPRMEYHVQTIEMQ-PAQIEGDLVLPPEVG 124
 QY 133 AAMQWLWRDAGIRACRYERRRREFHLLDSAVYLSHLERITEEGYVPTAQDVLRSRMPPTGI 192
 DB 125 GALEALWKDHGVQCFKRSREYQLNDSARYFDNIVRIATPDYMPNDQDVLRSRVKTGTGI 184
 QY 193 NEYCFVQKTNLRIVDVGQKSERKKWHCFENVIALIYLASLSYDQCLLENNQENRMK 252
 DB 185 TETTFIIGDLTYRMDFVGGQSRERKKWHCFENVTILFLVAISEYDQCLLENNQENRMK 244
 QY 253 ESALFQITILELPWFKSTSVILFLNKTDLIEEKIPTSHLATYFPSPGPKODAAAKRFI 312
 DB 245 EALTLDISICNSRWFKITSILFLNKTDLIEEKIPTSHLATYFPSPGPKODAAAKRFI 303
 QY 313 LDMYTRMYTCVDPGEGSKGARSRRFLSHVTCATDTQNIKRVKFDVRSVLARYL 368
 DB 304 LNRFEV-----SLNQHETKQIYTHETCATDTQIRFVMAAVNDIIQENL 347

RESULT 13

ID Q9HEP9 PRELIMINARY; PRT; 353 AA.
 AC Q9HEP9;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE G PROTEIN ALPHA SUBUNIT.
 GN G1.
 OS Botrytis cinerea (Botryotinia fuckeliana).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes;
 OC Helotiales; Sclerotiniaceae; Botryotinia.
 OX NCBI_TaxID=40559;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SAS56;
 RA Kasulke D., Tudzynski P., Tudzynski B.;
 RT "Cloning and characterisation of genes coding for G-alpha subunits
 from *Botrytis cinerea*.";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Y18436; CAC19871.1; -;
 SQ SEQUENCE 353 AA; 40983 MW; 01746CEID25ABFAF CRC64;

Query Match 38.5%; Score 755.5; DB 3; Length 353;
 Best Local Similarity 42.4%; Pred. No. 1.2e-54;
 Matches 151; Conservative 69; Mismatches 119; Indels 17; Gaps 5;

QY 15 TEDEKAARVDQENRILLQKQKDRGELKLLLLGPGESGKSTFIKQMRHHGAGYSEEE 74
 DB 7 TEEKEGAR-NEETENQKRDQKMMORNEIKMLLLGAGESGKSTILKQMKLHIEGGYSRDE 65
 QY 75 RKGFRLVYQNIQFVSMRAMIEMERLQIPFSRPSKHHASIVMSQDPYK--TTFEKRYA 132
 DB 66 RESFKEIFSNVQSMRVLAEESLEPLDQDQRAEYHVQTIEMQ-POQIEGDLNLPPEVG 124

QY 133 AAMQWLWRDAGIRACRYERRRREFHLLDSAVYLSHLERITEEGYVPTAQDVLRSRMPPTGI 192
 DB 125 SAIALWKDHGVQCFKRSREYQLNDSARYFDNIERIAQHDYMPNDQDVLRSRVKTGTGI 184
 QY 193 NEYCFVQKTNLRIVDVGQKSERKKWHCFENVIALIYLASLSYDQCLLENNQENRMK 252
 DB 185 TETTFIIGDLTYRMDFVGGQSRERKKWHCFENVTILFLVAISEYDQCLLENNQENRMK 244
 QY 253 ESALFQITILELPWFKSTSVILFLNKTDLIEEKIPTSHLATYFPSPGPKODAAAKRFI 312
 DB 245 EALTLDISICNSRWFKITSILFLNKTDLIEEKIPTSHLATYFPSPGPKODAAAKRFI 303
 QY 313 LDMYTRMYTCVDPGEGSKGARSRRFLSHVTCATDTQNIKRVKFDVRSVLARYL 368
 DB 304 LNRFEV-----SLNQHETKQIYTHETCATDTQIRFVMAAVNDIIQENL 347

RESULT 14

ID Q9NL97 PRELIMINARY; PRT; 357 AA.
 AC Q9NL97;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE G PROTEIN ALPHA SUBUNIT.
 GN HRG-ALPHA-N.
 OS Halocynthia roretzi (Sea squirt).
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
 OC Stolidobranchia; Pyruidae; Halocynthia.
 OX NCBI_TaxID=7729;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Iwasa T., Ohkuma M., Azuma T., Kanehara K., Watari A., Nakagawa M.,
 RA Kikkawa S., Tsuda M.;
 RT "A novel G protein alpha subunit gene expressed in the ectodermal
 lineage cells of an ascidian larva.";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB022098; BAA93630.1; -;
 DR InterPro: IPR001019; -;
 DR Pfam: PF00503; G-alpha; 1.
 DR PRINTS: PR00318; GPROTEINA.
 DR SMART: SM00275; G-alpha; 1.
 DR SMART: SM00275; G-alpha; 1.
 SQ SEQUENCE 357 AA; 40533 MW; 392901F92C3D519 CRC64;

Query Match 38.5%; Score 754; DB 5; Length 357;
 Best Local Similarity 43.3%; Pred. No. 1.6e-54;
 Matches 161; Conservative 66; Mismatches 115; Indels 30; Gaps 6;

QY 9 CCPWCLTEDEKAAARVDQENRILLQKQKDRGELKLLLLGPGESGKSTFIKQMRHHG 68
 DB 3 CAP-SKSENDKDAYSKSKEIDKQKDAENARKEVKLLLLGAGESGKSTIAQMKILHQD 61
 QY 69 GYSEERKGRPLVYQNIQFVSMRAMIEMERLQIPFSRPSKHHASIVMSQDPYK---- 124
 DB 62 GFSEERKGRFPVYVYTTNTIQSWAIVKAMESLIGDIYEDNERE-----EDGRKIRAS 113
 QY 125 TTFEK-----RYAAMQWLWRDAGIRACRYERRRREFHLLDSAVYLSHLERITEEGYVPTA 179
 DB 114 QTMEEIDITADIGALKRLKMDKGVACVARSREYQLNDSARYLALDRLCSPPDIITE 173
 QY 180 QDVLRSRMPPTGIINCYCFVQKTNLRIVDVGQKSERKKWHCFENVIALIYLASLSYD 239
 DB 174 QDVLTRVKTGTGIETTFQYKLNFTLIDVGQSRERKKWHCFQDVTAILFCVAMSAYD 233
 QY 240 QCLLENNQENRMKESLALFGTILELPWFKSTSVILFLNKTDLIEEKIPTSHLATYFPSPQ 299
 DB 234 QVLADEDETNRMIESLKLQFQICNNPFFAKTSMILFLNKKDLFEKIKKSPINTCFKEYE 293
 QY 300 GPKODAAAKRFILDMYTRMYTCVDPGEGSKGARSRRFLSHVTCATDTQNIKRVKFDV 359
 DB 294 GENSADASEH-IKEQF-----EAAKNKNCMEKIYTHETCATDTGNNRVFQAV 341

```
QY 360 RDSVLARYLDEI 371
- P78705
DB 342 SDVLMRKIIDNV 353

RESULT 15
ID P78705 PRELIMINARY; PRT; 353 AA.
AC P78705;
DT 01-MAY-1997 (TRENBLrel. 03, Created)
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE G PROTEIN ALPHA SUBUNIT.
GN GNA-1
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A;
RA Ivey F.D., Hodge P.N., Turner G.E., Borkovich K.A.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U56090; AAB37244.1; -.
DR HSP; P04896; IAZT.
DR InterPro; IPR001019; -.
DR Pfam; PF00503; G-alpha; 1.
DR PRINTS; PR00318; GPROTEINA.
DR SMART; SM00275; G-alpha; 1.
DR SMART; SM00275; G-alpha; 1.
SQ SEQUENCE 353 AA; 40729 MW; 87A376F964DA72E1 CRC64;

Query Match 38.4%; Score 753.5; DB 3; Length 353;
Best Local Similarity 42.4%; Pred. No. 1.7e-54;
Matches 151; Conservative 69; Mismatches 119; Indels 17; Gaps 5;

QY 15 TEDEKAAARVDQINILLEQKQDRGEKLLLLGPGSGKSTFIKQMLIHGAGYSEE 74
DB 7 TEEGEGAR-NEEENOLKRRDQQRNEIKMLLLGAGESGKSTILKMKLIHEGGYSDG 65
QY 75 RKGFRPLVYONIFVSMRAMIEAMERLQIPFSPESKHHASLVMSQDPYKV--TTFEKRYA 132
DB 66 AESFKEIIFSNVTQSMRVILEAMESLELFLADQQRVEYHVQTFPMQ--PAQIEGDVLPPEVG 124
QY 133 AAMQWLWRDAGIRACYRRERFHLDSAVYILSHLERITEEGYVPTAQDVLRSRMPPTGI 192
DB 125 NATEALWRDAGVQSCPKRSREYQLNDSARYYFDNIARIAAPDYPMPNDQDVLRSRVKTTGI 184
QY 193 NEYCFSVQKTNLRIVDVGQKSKRWIKHCFENVIALIVLASLEYDQCLEENNOENRMK 252
DB 185 TETTFITIGDLTVRMFDVGQSGRERKRWIKHCFENVTTILFLVAISEYDQLLFEDETVMRMQ 244
QY 253 ESLALFGTILELPWPKSTSVILFLNKTDLLEKIPTSHLATYFPSPGPKQDAEAAKRFI 312
DB 245 EALTLPDSICNSRWFTKTSILFLNKRDKFKELPVPSPMKNYFPDYEG--GDDYAAACDYI 303
QY 313 LDMYTRMYTCVDGPGSGKKGARSRLFSHYTCATDTQNIKRVFKDVRDVLARYL 368
DB 304 LARFV-----SLNQHETKQIYTHFTCAIDTTQIRFVMAAVNDIIQENL 347
```

Search completed: September 6, 2001, 10:55:04
Job time: 84 sec

Score	Strd Orig	ZScore	Escore	Len	Documentation
gb_pat1:AR095766	+ 1960.00	3223.09	7.5e-172	1125	! AR095766 Sequence 1 from paten
gb_pat1:AR0106753	+ 1960.00	3226.67	1.7e-171	2060	! AR0106753 Sequence 1 from paten
gb_pr10:HUMGAL16	+ 1960.00	3226.67	1.7e-171	2060	! M63904 Human G-alpha 16 protei
gb_pr10:AB015308	+ 1722.00	2834.05	1.3e-149	1766	! AB015308 Rattus norvegicus mRN
gb_om:AF169627	+ 1711.00	2820.61	7.1e-149	1125	! AF169627 Orcytolagus cuniculus
gb_pat1:AR095767	+ 1705.00	2810.67	2.6e-148	1125	! AR095767 Sequence 3 from paten
gb_pr10:MUSGNA15A	+ 1705.00	2808.71	3.3e-148	1353	! M80632 Mouse G protein alpha s
gb_pr10:BC005439	+ 1705.00	2804.70	5.5e-148	1974	! BC005439 Mus musculus, guanine
gb_ov:XLUU10502	+ 1057.00	1736.99	1.6e-88	1150	! U10502 Xenopus laevis guanine
gb_ov:XELGALPBIND	+ 1052.00	1727.03	5.8e-88	1347	! L05540 xenopus laevis alpha s
gb_om:DOGGA1P	+ 1047.00	1721.12	1.2e-87	1077	! L76257 Canis familiaris G alph
gb_om:PRF5AF2329284	+ 1047.00	1721.09	1.3e-87	1080	! AF2329284 Homo sapiens G alpha
gb_om:AF234260	+ 1046.00	1719.43	1.5e-87	1080	! AF234260 Rattus norvegicus het
gb_om:BOVG12A	+ 1046.00	1716.74	2.2e-87	1391	! D90336 Bovine mRNA for GRP-bir
gb_pr10:HSU43083	+ 1045.00	1712.96	3.5e-87	1700	! AF43083 Human G alpha-q (Gaq) m
gb_pr4:AF011496	+ 1044.00	1716.12	2.4e-87	1080	! AF011496 Homo sapiens GTP-bin
gb_pr10:HSU40038	+ 1044.00	1712.99	3.5e-87	1450	! U40038 Human GTP-binding prote
gb_pr10:MUSGAQ	+ 1044.00	1712.92	3.6e-87	1460	! M55412 Mouse G alpha q subunit
gb_pr10:HUMGALP	+ 1042.00	1712.83	3.6e-87	1077	! L76256 Homo sapiens G alpha q
gb_pr10:MUSGALL1	+ 1038.00	1703.70	1.2e-86	1364	! M55411 Mouse G alpha II subuni
gb_in3:HAUB9139	+ 1036.00	1659.73	1.9e-86	1451	! U89139 Homarus americanus hete
gb_ov:MGCA11A	+ 1035.00	1659.60	2.0e-86	1257	! F73072 M.gallopavo G-alpha-11
gb_pat1:AR073361	+ 1033.00	1697.90	2.4e-86	1080	! AR073361 Sequence 1 from patn
gb_pr4:AF011497	+ 1033.00	1697.90	2.4e-86	1080	! AF011497 Homo sapiens guanine
gb_pr10:HUMGTPBPRA	+ 1031.00	1690.82	6.1e-86	1540	! M60113 Human guanine nucleoti
gb_in2:AF201328	+ 1031.00	1689.77	6.9e-86	1700	! AF201328 Panulirus argus Gq/Il
gb_pr10:AF239674	+ 1030.00	1692.93	4.6e-86	1080	! AF239674 Rattus norvegicus gua
gb_ov:XLUU10494	+ 1026.50	1686.56	1.0e-85	1139	! U10494 Xenopus laevis guanine
gb_in3:DMW31092	+ 1026.00	1673.72	5.4e-85	3531	! U31092 Drosophila melanogaster
gb_in4:SODGCALF	+ 1017.50	1669.04	9.9e-85	1457	! L10289 Loligo forbesi visual G
gb_ov:AF059182	+ 1013.00	1662.63	2.3e-84	1320	! AF059182 Xenopus laevis G prot
gb_pr10:MUSGNA14A	+ 1013.00	1661.26	2.7e-84	1503	! M80631 Mouse G protein alpha s
gb_om:BOVG11A	+ 1013.00	1660.13	3.1e-84	1671	! D90335 Bovine mRNA for GRP-bir
gb_in1:AB025782	+ 1006.00	1650.71	1.0e-83	1362	! AB025782 Octopus vulgaris OvGa
gb_pr4:AF105201	+ 1005.00	1646.96	1.7e-83	1658	! AF105201 Homo sapiens G-protei
gb_in3:CVI250443	+ 1001.00	1641.91	3.2e-83	1429	! AJ250443 Calliphora vicina mRN
gb_in3:LSGPASQG	+ 995.00	1631.57	1.2e-82	1484	! Z23106 L.stagnalis mRNA for G
gb_in1:APO08456	+ 989.50	1623.41	3.4e-82	1357	! APO08456 ratinopecten yessoensis
gb_in3:LPU08586	+ 986.00	1615.23	9.8e-82	1698	! UB8586 Limulus polyphemus Gq p
gb_in2:AY008139	+ 974.00	1600.28	6.7e-81	1068	! AY008139 Caenorhabditis elegans

```

151 ArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLe 167
|||||
451 CGCGGGAATTCACCTGCTCGATTACGCGGTGACTACCTGTCCACCT 500
|||||
167 uGluArgIleThrGluGluGlyTyrValProThrAlaGlnAspValLeuA 184
|||||
501 GGAGCGCATCACCGAGAGGGCTACGTCACACAGCTCAGGACGTGCTCC 550
|||||
184 rGSerArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGln 200
|||||
551 GCAGCCGATGCCACCATCGGCATCAACAGTACTGCTTCTCGGTCCAG 600
|||||
201 LysThrAsnLeuArgIleValAspValGlyGlnLysSerGluArgLy 217
|||||
601 AAAACCACTGCGGATCGTGGAGCTCGGGGCCAGGAGCAGAGCGTAA 650
|||||
217 sLysTrpIleHisCysPheGluAsnValIleAlaLeuIleTyrLeuAla 234
|||||
651 GAAATGGATCCATTGTTTCGAGAAGCTGATCGGCCCTCATCTACCTGGCCT 700
|||||
234 erLeuSerGluTyrAspGlnCysLeuGluGluAsnAsnGlnGluAsnArg 250
|||||
701 CACTGAGTGAATACGACCACTGCTGGAGAGAACAAACAGGAGAACCGC 750
|||||
251 MetLysGluSerLeuAlaLeuPheGlyThrIleLeuGluLeuProTrpPh 267
|||||
751 ATGAAGGAGAGCCTCGCATTTGTTGGGACTATCTCGGAACCTACCTGGTT 800
|||||
267 eLysSerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuGluG 284
|||||
801 CAAAGCACATCCGTCATCTCTCTTCACAAAACCCGACATCTCGGAGG 850
|||||
284 lLysIleProThrSerHisLeuAlaThrTyrPheProSerPheGlnGly 300
|||||
851 AGAAATCCCCACCTCCACCTGGCTACTATTTCCTCCAGCTTCCAGGGC 900
|||||
301 ProLysGlnAspAlaGluAlaAlaLysArgPheIleLeuAspMetTyrTh 317
|||||
901 CCTAAGCAGATGCTGAGGAGGAGCAAGAGTTTCATCTCGACATGTACAC 950
|||||
317 rArgMetTyrThrGlyCysValaspGlyProGluGlySerLysLysGlyA 334
|||||
951 GAGATGTACACCGGGTGGTGGAGCGCCCGGAGGCGAGCAAGAAGGGCG 1000
|||||
334 laArgSerArgArgLeuPheSerHisTyrThrCysAlaThrAspThrGln 350
|||||
1001 CAGATCCCGACGCTTTTCAGCACCTACACATGTGCCACAGACACAG 1050
|||||
351 AsnIleArgLysValPheLysAspValArgAspSerValLeuAlaArgTy 367
|||||
1051 AACATCCGCAAGGTCTTCAAGGAGCTGCGGGAGCTCGGTGTGCTCGCCGCTA 1100
|||||
367 rLeuAspGluIleAsnLeuLeu 374
|||||
1101 CCTGGACGAGATCAACTGCTG 1122

```

seq_name: gb_pat1:AR106753

```

seq_documentation_block:
LOCUS AR106753 2060 bp DNA PAT 14-FEB-2001
DEFINITION Sequence 1 from patent US 6107091.
ACCESSION AR106753
VERSION AR106753.1 GI:12821283
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 2060)
AUTHORS Cowser, L.M.
TITLE Antisense inhibition of G-alpha-16 expression
JOURNAL Patent: US 6107091-A 1 22-AUG-2000;
FEATURES
Location/Qualifiers

```

```

source 1..2060
BASE COUNT 415 a 631 c 640 g 374 t
ORIGIN

alignment_scores:
Quality: 1960.00 length: 374
Ratio: 5.241 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-878-801-2 x AR106753
Align seg 1/1 to: AR106753 from: 1 to: 2060

1 MetAlaArgSerLeuThrTrpArgCysCysProTrpCysLeuThrGluAs 17
|||||
220 ATGGCCCGCTCGCTGACCTGGCGCTGCTGCCCTGGGTGCCTGACGGAGGA 269
|||||
17 pGluLysAlaAlaAlaArgValaspGlnGluIleAsnArgIleLeuLeuG 34
|||||
270 TGAGAAAGCGCCGCGCGGGTGGACACAGAGATCAACAGGATCCTCTTGG 319
|||||
34 luGlnLysLysGlnAspArgGlyGluLeuLysLeuLeuLeuGlyPro 50
|||||
320 AGCAGAAGAAGCAGGACCGCGGGGAGCTGAAG:TCGCTGCTTTGGGCCCA 369
|||||
51 GlyGluSerGlyLysSerThrPheIleLysGlnMetArgIleIleHisG 67
|||||
370 GCGGAGAGCGGAAGAGACACTTCATCAAGCA:ATGCGGATCATCCACGG 419
|||||
67 yAlaGlyTyrSerGluGluArgLysGlyPheArgProLeuValTyrG 84
|||||
420 CGCGCGCTACTCGGAGGAGGCGCAGGGCTTCCGGCCCTGGTCTACC 469
|||||
84 InasnIlePheValSerMetArgAlaMetIleIleAlaMetGluArgLeu 100
|||||
470 AGAACATCTTCGTGCTCCATCGGGCCATGATCAGGCCATGGAGCGGCTG 519
|||||
101 GlnIleProPheSerArgProGluSerLysHisHisAlaSerLeuValMe 117
|||||
520 CAGATTCATTCAGACGCGCCGAGAGCAAGCA:CAACGCTAGCCTGGTGCAT 569
|||||
117 tSerGlnAspProTyrLysValThrPheGluLysArgTyrAlaAlaA 134
|||||
570 GAGCCAGGACCCCTATAAAGTGACACGTTTGAGAAGCGCTACGCTCGG 619
|||||
134 laMetGlnTrpLeuTrpArgAspAlaGlyIleArgAlaCysTyrGluArg 150
|||||
620 CCATGCAAGTGGCTGTGGAGGAGTCCCGGCATC:GGGCTGCTATGAGCGT 669
|||||
151 ArgArgGluPheHisLeuLeuAspSerAlaVal:TyrtTyrLeuSerHisLe 167
|||||
670 CGCGGGGAATTCACCTGCTCGATTTCAGCCGT:TTACTACCTGTCCCACT 719
|||||
167 uGluArgIleThrGluGluGlyTyrValProThrAlaGlnAspValLeuA 184
|||||
720 GGAGCGCATCACCGAGGAGGCTACGTCGCCCA:AGCTCAGGACGTGCTCC 769
|||||
184 rGSerArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGln 200
|||||
770 GCAGCCGATGCCACCATCGGCATCAACAG:ACTGCTTCTCCGTCCAG 819
|||||
201 LysThrAsnLeuArgIleValaspValGlyGlnLysSerGluArgLy 217
|||||
820 AAAACCAACTGCGGATCGTGGAGCTCGGGGCCAGAGTCAAGAGTCAGAGCGTAA 869
|||||
217 sLysTrpIleHisCysPheGluAsnValIleAlaLeuIleTyrLeuAla 234
|||||
870 GAAATGGATCCATTGTTTCGAGAAGCAGTACG:CTCATCTACCTGGCCT 919
|||||
234 erLeuSerGluTyrAspGlnCysLeuGluGluAsnGlnGluAsnArg 250

```

```
|||||
920 CACTGAGTGAATACGACCACTGCTGGAGGAGAACCAACGAGGAGACCGC 969
251 MetLysSerLeuAlaLeuPheGlyThrIleLeuGluLeuProThrPh 267
|||||
970 ATGAAGAGAGCCCTCCGATGTTGGGACTATCTCTGGAACCTACCTGGTT 1019
267 eLysSerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuGluG 284
|||||
1020 CAAAAGCACATCGCTCATCTCTTCTCAACAAACCGACATCTCTGGAGG 1069
284 LuLysIleProThrSerHisLeuAlaThrThrPheProSerPheGlnGly 300
|||||
1070 AGAAATCCCACTCCACCTCCGCTGCTACCTATTTCCCACTTCCAGGGC 1119
301 ProLysGlnAspAlaGluAlaLysArgPheIleLeuAspMetTyrTh 317
|||||
1120 CCTAAGCAGGATGCTGAGGAGCAAGAGGTTTCATCTGGACATGTACAC 1169
317 rArgMetTyrThrGlyCysValAspGlyProGluGlySerLysLysGlyA 334
|||||
1170 GAGGATGTACACCGGCTGCTGACGCGCCGAGGCGCAAGAGGGCG 1219
334 laArgSerArgLeuPheSerHisTyrThrCysAlaThrAspThrGln 350
|||||
1220 CACGATCCGACGCGCTTTTCCAGCCACTACACATGTGCCACACACAG 1269
351 AsnIleArgLysValPheLysAspValArgAspSerValLeuAlaArgTy 367
1270 AACATCGCAAGGCTTCAAGAGAGCTGCGGAGCTCGGTGCTGCCCGCTA 1319
367 rLeuAspGluIleAsnLeuLeu 374
|||||
1320 CCTGGACGAGATCAACCTGCTG 1341
seq_name: gb_pr10:HUMGA16

seq_documentation_block: 2060 bp. mRNA PRI 27-APR-1993
LOCUS HUMGA16
DEFINITION Human G-alpha 16 protein mRNA, complete cds.
ACCESSION M63904
VERSION M63904.1 GI:182891
KEYWORDS G-alpha 16 protein.
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
REFERENCE
AUTHORS Anatruda,T.I.II., Steele,D.A., Slepak,V.Z. and Simon,M.I..
TITLE G-alpha16, a G protein alpha subunit specifically expressed in
hematopoietic cells
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 88, 5587-5591 (1991)
MEDLINE 91288509
FEATURES
source Location/Qualifiers
1..2060
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HL-60"
<1..2060
/product="G-alpha-16 protein"
220..1344
/codon_start=1
/product="G-alpha-16 protein"
/protein_id="AAA35860.1"
/db_xref="GI:182892"
/translation="MARSLTWRCPCWCLTEDEKAAARVDOEINRILLEOKKODRGELK
LLLGPGESKSTFIKQMLIHGAGYSEERKGRFLYVQNIQFVSMRAIEAMERLQI
PFSRPSKHSKLVMSQDPYKVTTFKKRYAAAMQWLRDAGRACRYRERPFLLDSA
VYLSHLERITEGYVPTAQDLVRSMPTTGINCYCFVQKTNLRIVDVGGQSKERK
WIHCFENVIALIVLASLSEYDOCLENNQENRKNESIALFGTILELPWFKSTSVILFL
NKTDILEEKIPTSLATYFPFSGPKODAEAAKRFILDMYTRMTYTCVDGPEGSKGA
RSRLFSHYTCARDTDQNIKRVKFDVRDVSILRYLDEINLL"
BASE COUNT 415 a 631 c 640 g 374 t
```

```
ORIGIN
alignment_scores:
Quality: 1960.00 Length: 374
Ratio: 5.241 Gaps:
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-08-878-801-2 x HUMGA16
Align seg 1/1 to: HUMGA16 from: 1 to: 2060
1 MetAlaArgSerLeuThrTrpArgCysCysProTrpCysLeuThrGluAs 17
220 ATGGCCCGCTCGCTGACCTGGCGCTGCTGCCCTGGTGCCTGACGAGGA 269
17 pGluLysAlaAlaAlaArgValAspGlnGluIleAsnArgIleLeuLeu 34
270 TGAGAAGCGCGCGCGCGGTGGACGAGATCAACAGGATCCTCTTGG 319
34 LuGlnLysLysGlnAspArgGlyGluLeuLysLeuLeuLeuGlyPro 50
320 ACAGAAGACGAGGACCGCGGGAGCTGAAGCTGCTGCTTTGGGCCCA 369
51 GlyGluSerGlyLysSerThrPheIleLysGlnMetArgIleIleHisG 67
370 GCGAGAGCGGGAAGACACCTTCATCAAGCAGATCGGATCATCCACGG 419
67 yAlaGlyTyrSerGluGluArgLysGlyPheArgProLeuValTyrG 84
420 CGCGCGCTACTCGGAGGAGGAGCGCAAGGCTTCGCGCCCTGGTCTACC 469
84 InAsnIlePheValSerMetArgAlaMetIleGluAlaMetGluArgLeu 100
470 AGAACATCTTCGTGTCATCGCGGCATCATCGAGGCCATGAGGCGGTG 519
101 GlnIleProPheSerArgProGluSerLysHisHisAlaSerLeuValMe 117
520 CAGATTCCATTTCAGCAGCGCCGAGCAAGCACCACGCTAGCCTGGTCA 569
117 tSerGlnAspProTyrLysValThrThrPheGluLysArgTyrAlaAla 134
570 GAGCCAGGACCCCTATAAAGTGACACAGCTTGAGAAAGCGCTACGCTG 619
134 laMetGlnTrpLeuTrpArgAspAlaGlyIleArgAlaCysTyrGluArg 150
620 CCATGCAGTGGCTGTGGAGGATCGCGGCATCGCGGCCTGCTATGAGCG 669
151 ArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLe 167
670 CGGCGGGAATTCACCTGCTCGATTGAGCCGTGCTACTACTCTGCCACCT 719
167 uGluArgIleThrGluGlyTyrValProThrAlaGlnAspValLeuA 184
720 GGAGCGCATCCCGAGGAGGCTACGTCCCCACAGCTCAGGACGTGCTCC 769
184 rgSerArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGln 200
770 GCAGCGCATGCCACCTGCGCATCAACGAGTACTGCTTCTCCGTGCG 819
201 LysThrAsnLeuArgIleValAspValGlyGlnLysSerGluArgLy 217
820 AAAACCAACCTGCGGATCGGTGGACGCTCGGGGGCCAGAGTCAGAGCTAA 869
217 sLysTrpIleHisCysPheGluAsnValIleAlaLeuIleTyrLeuAla 234
870 GAATGGATCCATGTTTTCAGAAACGTGATCGCCCTCATCTACTGCGCT 919
234 erLeuSerGluTyrAspGlnCysLeuGluGluAsnAsnGlnGluAsnArg 250
920 CACTGAGTGAATACGACAGTGCCTGGAGGAGAGAACACAGGAGAACCCG 969
```

251 MetLysGluSerLeuAlaLeuPheGlyThrIleLeuGluLeuProTrpPh 267
 |||||
 970 ATGAGGAGAGGCTCCATGTTGGACTATCTTGGAACTACCCGGTT 1019
 267 eLysSerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuGluG 284
 |||||
 1020 CAAAGACATCCGTCATCTCTTCTCAACAAACCCGACATCCTGGAGG 1069
 284 LuLysIleProThrSerHisLeuAlaThrTyrPheProSerPheGlnGly 300
 |||||
 1070 AGAAATCCCACTCCCACTCCGCTGCTATCTTCCAGTTTCCAGGGC 1119
 301 ProLysGlnAspAlaGluAlaLysArgPheIleLeuAspMetTyrTh 317
 |||||
 1120 CTAAGCAGGATGCTGAGCAGCAAGAGTTTCATCTGGACATGATAC 1169
 317 rArgMetTyrThrGlyCysValAspGlyProGluGlySerLysGlyA 334
 |||||
 1170 GAGGATGATACCGGGTGGTGGAGCGCCCGAGGAGCAAGAGGGC 1219
 334 LaArgSerArgArgLeuPheSerHisTyrThrCysAlaThrAspThrGln 350
 |||||
 1220 CACGATCCGAGCGCTTTCAGCCACTACATGTTGCCACAGACACAG 1269
 351 AsnIleArgLysValPheLysAspValArgAspSerValLeuAlaArgTy 367
 |||||
 1270 AACATCCCAAGGTTCTCAAGGACGTGGGGACTCGGTGCTCGCCCGCTA 1319
 367 rIeuAspGluIleAsnLeuLeu 374
 |||||
 1320 CTGGACGAGATCAACCTGCTG 1341
 seq_name: gb_rol:AB015308

seq_documentation_block:
 LOCUS AB015308 1766 bp mRNA ROD 06-FEB-1999
 DEFINITION Rattus norvegicus mRNA for GTP binding protein alpha 15, complete cds.

ACCESSION AB015308
 VERSION AB015308.1 GI:3219264
 KEYWORDS GTP binding protein alpha 15.
 SOURCE Rattus norvegicus tissue_lib:tongue epithelium cDNA to mRNA.
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 1766)
 AUTHORS Kusakabe,Y. and Abe,K.
 TITLE Direct Submission
 JOURNAL Submitted (05-JUN-1998) to the DDBJ/EMBL/GenBank databases. Yoko Kusakabe, University of Tokyo, Dept. Applied Biol. Chem.; 1-1-1, Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan (E-mail:aka7308@hongo.ecc.u-tokyo.ac.jp, Tel:81-3-5802-8897, Fax:81-3-5802-8897)

REFERENCE 2 (sites)
 AUTHORS Kusakabe,Y., Yamaguchi,E., Tanemura,K., Kameyama,K., Chiba,N., Arai,S., Emori,Y. and Abe,K.
 TITLE Identification of two alpha-subunit species of GTP-binding proteins, Galphais and Galphag, expressed in rat taste buds
 JOURNAL Biochem. Biophys. Acta 1403 (3), 265-272 (1998)
 MEDLINE 98351893

FEATURES
 Location/Qualifiers
 1..1766

source
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 /tissue_lib="tongue epithelium"
 231..1355

CDS
 /codon_start=1
 /product="GTP binding protein alpha 15"
 /protein_id="BAA28827.1"
 /db_xref="GI:3219265"
 /translation="MARSLTWGCCPWCILTEFEKTAARDIOEINKLLEOKQBERGELK
 LLLLPGESGSKSTFIKMRRIIHGAGYSEEDRAFRLLVYQNFVSMQAWTEAMDRLQI

PFSPDQKSHASLVMTQDPYKVSSFEKPYAVAMOYLWRDAGIRACIYERRRREFHLLDSA
 VYLSHLERIAEDDYIPTAQDVLRSRMPTTGINNEYCFSVOKTKLRIVDAGGOKSERKK
 WICFENVIALIYLASLSEV)OCLEENSOENRMKESLAFSTILELPWFKSTSVILEL
 NKTDILEDKIHITSHLASVFP:FOGPRRDAAAKRFILDMYARVYASCAEPHDGGRKGS
 RARLFAHETCATDTHSVRS:/FKDVRDSVLARYLDEINLL"
 1766

polyA_site

BASE COUNT 401 a 516 c 513 g 336 t
 ORIGIN

alignment_scores:
 Quality: 1722.00 length: 374
 Ratio: 4.692 gaps: 0
 Percent Similarity: 98.128 Percent Identity: 86.096

alignment_block:

US-08-878-801-2 x AB015308 ..

Align seg 1/1 to: AB015308 from: 1 to: 1766

1 MetAlaArgSerLeuThrTrpArgCysCysProTipCysLeuThrGluAs 17
 |||||
 231 ATGGCCCGTCCCTGACTTGGGGCTGCTGCTCCCTGGCTGACGGAAGA 280
 17 pEluLysAlaAlaAlaArgValAspGlnGluIleAsnArgIleLeuLeuG 34
 |||||
 281 GGAGAAGACTGCGCCAGACATCGACCAGGAGA1CAACAAGATTTTTGTGG 330
 34 LuGlnLysLysGlnAspArgGlyGluLeuLysIleuLeuLeuGlyPro 50
 |||||
 331 AACAGAAGAAACAGAGCGCGGGGAATTGAAAC TCCTGCTGTTGGGGCCC 380
 51 GlyGluSerGlyLysSerThrPheIleLysGlnMetArgIleIleHisG 67
 |||||
 381 GGTGAGACGGGAAAGACACGTTTCATCAAGCAPATGCCATCATTCACGG 430
 67 ValaGlyTyrSerGluGluGluArgLysGlyPheArgProLeuValTyrG 84
 |||||
 431 CGCCGGCTACTCTGAGGAGGACCGAGAGCCTTCGGCTGCTGCTACCC 480
 84 InAsnIlePheValSerMetArgAlaMetIleGluAlaMetGluArgLeu 100
 |||||
 481 AGAACATCTTCGTCTCCATGAGGCCCATGATTGAAGCAATGACAGCGCTG 530
 101 GlnIleProPheSerArgProGluSerLysHisAlaSerLeuValme 117
 |||||
 531 CAGATCCCTTCAGACGGCGGACGAGCAAAACAGACGCCAGCCTGGTGAT 580
 117 tSerGlnAspProTyrLysValThrPheGlyLysArgTyrAlaAla 134
 |||||
 581 GACCCAGGACCCCTATAAAGTGAGCTCGTTCGA3AAGCCATATGACGTGG 630
 134 laMetGlnTyrLeuTyrArgAspAlaGlyIleArgAlaCysTyrGluArg 150
 |||||
 631 CCATGCAGTACCTGTGGCGGACGCGGGCATCC3CGCATGCTACGACGG 680
 151 ArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLe 167
 |||||
 681 AGCGGTGAATTCACCTGCTGGACTCCCGCGTGACTACCTGTACACACCT 730
 167 uGluArgIleThrGluGluGlyTyrValProTh::AlaGlnAspValLeu 184
 |||||
 731 GGAGCGCATCCCGAGGACGACACTACATCCCCAC::GGCAGGAGCGTGCTG 780
 184 rgSerArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGln 200
 |||||
 781 GCAGTCGATGCCACGACCATGGCATCAATGAGTACTGCTTTTCCGTGCAG 830
 201 LysThrAsnLeuArgIleValAspValGlyGlyLysSerGluArgLy 217
 |||||
 831 AAAACCAAACTGCGTCATCGTGGATGCTGGCGGC3AGAAGTCAGAACGTAA 880

```
217 sLysTrpIleHisCysPheGluAsnValIleAlaLeuIleTyrLeuAlaLas 234
|||||
881 GAAATGATCACTGTTTCGAGACGCTGATTGGCCCTCATCTACCTGGCGT 930
234 erLeuSerGluTyrAspGlnCysLeuGluGluAsnGlnGluAsnArg 250
|||||
931 CTCGACGAGTATGACCATGCTGCTGGAGAGACACATCAGGAGACCGT 980
251 MetLysGluSerLeuAlaLeuPheGlyThrIleLeuGluLeuProTrrpPh 267
|||||
981 ATGAAGAGAGTCTCGCTCTGTTTAGCACGATCCTAGAGCTGCCCTGTT 1030
267 eLysSerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuGluG 284
|||||
1031 CAAGAGCACTCGGTATCTCTCTCCACCAAGACACATCTCTGGAGG 1080
284 luLysIleProThrSerHisLeuAlaThrTyrPheProSerPheGlnGly 300
|||||
1081 ATAAATCCACACCTCCACCTAGCTCCTATCTTCCCGAGCTTCCAGGGA 1130
301 ProLysGlnAspAlaGluAlaAlaLysArgPheIleLeuAspMetTyrTh 317
|||||
1131 CCCCGAGGAGACGAGAGCGCGCAAGCGCTTCATCTTGGACATGTACGC 1180
317 rArgMetTyrThrGlyCysValAspGlyProGluGlySerLysLysGlyA 334
|||||
1181 GCGGTGTACCGAGCTGTGAGAGCCCCACGAGCGGTGGCAGAGGAT 1230
334 laArgSerArgArgLeuPheSerHisTyrThrCysAlaThrAspThrGln 350
|||||
1231 CCCGCGCGCGCGCTCTTCCACACATCTTCCCTGTGCGACGACACGCAC 1280
351 AsnIleArgLysValPheLysAspValArgAspSerValLeuAlaArgTy 367
1281 AGCGTCCCGACGCTGTCAAGAGCGTGGGGACTCAGTGTGGCCCGGTA 1330
367 rLeuAspGluIleAsnLeuLeu 374
|||||
1331 CTGGACGAGATCAACCTGTTG 1352

seq_name: gb_om:AF169627
seq_documentation_block:
LOCUS AF169627 1125 bp mRNA MAM 09-NOV-1999
DEFINITION Oryctolagus cuniculus Galphal6 mRNA, complete cds.
ACCESSION AF169627
VERSION AF169627.1 GI:6288791
KEYWORDS
SOURCE
ORGANISM
Oryctolagus cuniculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE
1 (bases 1 to 1125)
Feild,J.A., Foley,J.J., Testa,T.T., Nuthulaganti,P., Ellis,C.,
Sarau,H.M. and Ames,R.S.
Cloning and characterization of a rabbit ortholog of human Galphal6
and mouse Galphal6
FEBS Lett. 460 (1), 53-56 (1999)
20035837
2 (bases 1 to 1125)
Feild,J.A., Testa,T.T., Foley,J.J., Nuthulaganti,P., Ellis,C.,
Sarau,H.M. and Ames,R.S.
Direct Submission
Pharmaceuticals, 709 Swedeland Rd., King of Prussia, PA 19406-0939,
USA
FEATURES
source
1..1125
/organism="Oryctolagus cuniculus"
/db_xref="taxon:9986"
1..1125
/note="G-protein alpha subunit"
/codon_start=1
```

```
/product="Galphal6"
/protein_id="AAF06740.1"
/db_xref="GI:6288792"
/translation="MARSLAWRCPCWCLSEDEKAAARVDQEIETRLLEHRRVGRGELK
LLLLGTGSGKSTFIKQMRITIHGAGYSEDRKGFRPLVQNFISVQALIEAMDRLQI
PYSPESKLHSLVMSQDPYKVTETRYALAVQSLWRDAGVTRACIERRERHLLDSA
VYVLSHLERIAIEGVPTAODVLRSMPTTGNEYCFVQKTNLRLIVDVGQKSRKK
WTHCFEDTALIFLASLSEYDQGLENGENRQMSALFGLVLPWFRAVTSVILFL
NKTDILEDKVRTSHLATYPPGFRGPPQDPEAAKRFILLEYTRYVYAGAAAGPDGASKGP
RSRRLFSHYTCATDQNIKRVKFDVRSVLAARYLDEINLL"
BASE COUNT 207 a 386 c 364 g 168 t
ORIGIN

alignment_scores:
Quality: 1711.00 Length: 374
Ratio: 4.740 Gaps: 0
Percent Similarity: 96.524 Percent Identity: 85.829

alignment_block:
US-08-878-801-2 x AF169627 ..
Align seg 1/1 to: AF169627 from: 1 to: 1125

1 MetAlaArgSerLeuThrTrrpArgCysCysProTrrpCysLeuThrGluAs 17
|||||
1 ATGGCCCGCTCCCTGGCTGGCGCTGCTGCCCTGCTGCTGAGCGAGGA 50

17 pGluLysAlaAlaAlaArgValAspGlnGluIleAsnArgIleLeuLeuG 34
|||||
51 CGAGAAAGCGCGCGCTCGCTGCGACCAAGAAATCACCAGGCTGCTCTGG 100

34 luGlnLysLysGlnAspArgGlyGluLeuLysLeuLeuLeuLeuPro 50
|||||
101 AACACAGCGCGCAGGTGCGAGCGAGCTCAAACTGCTGCTGCTGGCGAG 150

51 GlyGluSerGlyLysSerThrPheIleLysGlnMetArgIleIleHisGl 67
|||||
151 GCGGAGAGCGGCAAGAGCACGTTTCATCAAGCAGATGCCCATCATCCACG 200

67 yAlaGlyTyrSerGluGluGluArgLysGlyPheArgProLeuValTyrG 84
|||||
201 CCGCGGCTACTCGAGAGGAGCGCAAGGCTTCGCGCGCTGCTCTTCC 250

84 InAsnIlePheValSerMetArgAlaMetIleGluAlaMetGluArgLeu 100
|||||
251 AGAACATCTTCTCTCGGTCCAGGCGCATCATCGAGGCGCATGGACCGCTG 300

101 GlnIleProPheSerArgProGluSerLysHisAlaSerLeuValMe 117
|||||
301 CAGATCCCCCTACAGCGCGCGAGAGCAAGCTCCAGCGCGCTGGTCTAT 350

117 tSerGlnAspProTyrLysValThrPheGluLysArgTyrAlaAla 134
|||||
351 GAGCCAGGAGCCCTACAGGTGAACACGTTCCGAGCGCGCTAGCCCTG 400

134 laMetGlnTrpLeuTrrpArgAspAlaGlyIleArgAlaCysTyrGluArg 150
|||||
401 CCGTGCAGAGCGCTGTGGAGGAGCGCGGCGCTCGGGCGCTGCTACGAGCG 450

151 ArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLe 167
|||||
451 CGCGGGAGTCTCACCTGCTGGACTCGGCGGTGCTACTACTGTCGCACCT 500

167 uGluArgIleThrGluGluGlyTyrValProThrAlaGlnAspValLeuA 184
|||||
501 CGAGCGCATCGCGAGAGGCTACGTGCCAGCGCGCAGGACGTGTTC 550

184 rgSerArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGln 200
|||||
551 GTAGCGCATGCCACACCGCATCAATGAGTACTGCTTCTCCGTCCGTCAG 600

201 LysThrAsnLeuArgIleValAspValGlyGlyGlnLysSerGluArgLy 217
```

```

|||||
601 AAGACCAACCTCGCATCGTGGACGTCGGGGCCAGAAATCGAGCGCAG 650
|||||
217 slystrpIleHisCysPheGluAsnValIleAlaLeuIleTyrLeuAlas 234
|||||
651 GAAGTGGATCCACTCTTCGAGGAGCTGACGGCGCTCATCTTCCTGGCCT 700
|||||
234 erLeuSerGluTyrAspGlnCysLeuGluGluAsnGlnGluAsnArg 250
|||||
701 CACTGAGCGAGTACGACAGTGCCTGGAGGAGACGGCCAGGAGAACCGC 750
|||||
251 MetLysGluSerLeuAlaLeuPheGlyThrIleLeuGluLeuProTrpPh 267
|||||
751 ATGACGAGGAGCGCTGGCGCTGTCGGCACCGCTCTGGCGCTGCGCTGTT 800
|||||
267 eLysSerThrValIleLeuPheLeuAsnLysThrAspIleLeuGluG 284
|||||
801 CCGGGCCACCTCGGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 850
|||||
284 LuLysIleProThrSerHisLeuAlaThrTyrPheProSerPheGlnGly 300
|||||
851 ACAAGGTCGCGACGTCGACCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
|||||
301 ProLysGlnAspAlaGluAlaLysArgPheIleLeuAspMetTyrTh 317
|||||
901 CCGCGGACGACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 950
|||||
317 rArgMetTyrThrGlyCysValAspGlyProGluGlySerLysLysGlyA 334
|||||
951 ACGCGTGTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1000
|||||
334 laArgSerArgArgLeuPheSerHisTyrThrCysAlaThrAspThrGln 350
|||||
1001 CGCGCTCCGCGCGCTCTTCAGGACCTACACGTGCGCGCACGACGAC 1050
|||||
351 AsnIleArgLysValPheLysAspValArgAspSerValLeuAlaArgTy 367
|||||
1051 AACATACGCAAGGTCTTCAAGGACGTGCGGACTCGTGTCTGCGCGCTA 1100
|||||
367 rLeuAspGluIleAsnLeuLeu 374
|||||
1101 CCTGGACGAGATCAACCTGCTG 1122

```

seq_name: gb_pat1:AR095767

```

seq_documentation_block: 1125 bp DNA PAT 08-SEP-2000
LOCUS AR095767
DEFINITION Sequence 3 from patent US 6004808.
ACCESSION AR095767
VERSION AR095767.1 GI:10023948
KEYWORDS
SOURCE unknown.
ORGANISM unknown.

```

```

REFERENCE
  1 (bases 1 to 1125)
  Negulescu, P., Offermanns, S., Simon, M. and Zuker, C.
  TITLE Promiscuous G-protein compositions and their use
  JOURNAL Patent: US 6004808-A 3 21-DEC-1999;
  FEATURES
    Location/Qualifiers
      1..1125
        /organism="unknown"

```

```

BASE COUNT 263 a 328 c 325 g 209 t
ORIGIN

```

```

alignment_scores:
  Quality: 1705.00
  Ratio: 4.658
  Percent Similarity: 97.861
  Percent Identity: 84.759

```

alignment_block:

US-08-878-801-2 x AR095767

Align seg 1/1 to: AR095767 from: 1 to: 1125

```

1 MetAlaArgSerLeuThrTrpArgCysProTyrCysLeuThrGluAs 17
|||||
1 ATGGCCCGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 50
|||||
17 pGluLysAlaAlaAlaArgValAspGlnGluLysLeuLeuLeuLeu 34
|||||
51 GGAGAGAGACTGCGCCAGATCGACCCAGAGAGAACTCAACAGGATTT 100
|||||
34 LuGlnLysLysGlnAspArgGlyGluLeuLysLeuLeuLeuLeuPro 50
|||||
101 AACAGAAAAACAAGAGCGCGAGGAAATTTGAAATCTCTGCTGCTGG 150
|||||
51 GlyGluSerGlyLysSerThrPheIleLysGlnMetArgIleIleHis 67
|||||
151 GGTGAGAGCGGAGAGTACGTTTCATCAAGCAATGCGCATCATTCAC 200
|||||
67 yAlaGlyTyrSerGluGluArgLysGlyPheArgProLeuValTyrG 84
|||||
201 TGTGGGCTACTCGGAGGAGGACCGCAGAGCGCTTCGGGCTGCTCAT 250
|||||
84 InAsnIlePheValSerMetArgAlaMetIleGluAlaMetGluArg 100
|||||
251 AGAACATCTTCTGCTCCATGCGAGCCATGATAGATGCGATGACGCG 300
|||||
101 GlnIleProPheSerArgProGluSerLysHisAlaSerLeuValMe 117
|||||
301 CAGATCCCTTTCAGCAGCGCTGACAGCAAGCAACAGCCACGCTAGT 350
|||||
117 tSerGluAspProTyrLysValThrPheGluLysArgTyrAlaAla 134
|||||
351 GACCCAGGACCCCTATTAAGTGAGCACAATTCGGAAGCCATATCAG 400
|||||
134 laMetGlnTrpLeuTrpArgAspAlaGlyIlePrgAlaCysTyrGlu 150
|||||
401 CCATGCACTACCTGTGCGGACGCGGCGCATCCGTGATGTACGAGCA 450
|||||
151 ArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHis 167
|||||
451 AGCGTGAATTCACCTTCTGGACTCCGCGGTGTATTACCTGTCCACCT 500
|||||
167 uGluArgIleThrGluGluGlyTyrValProThrAlaGlnAspVal 184
|||||
501 GGAGCGCATATCAGAGACAGACTACATCCCCACATCGCGAAGACGT 550
|||||
184 rgSerArgMetProThrThrGlyIleAsnGluTyrCysPheSerVal 200
|||||
551 GCAGTCGATGCGCCACACAGGATCAATGAGTACTGCTTCTCCGTGA 600
|||||
201 LysThrAsnLeuArgIleValAspValGlyGly3LysSerGluArg 217
|||||
601 AAAACCAAACTGCGCATCTGTTGGTGGCTAGAGGTACAGACGCTAG 650
|||||
217 sLysTrpIleHisCysPheGluAsnValIleAlaLeuIleTyrLeu 234
|||||
651 GAAATGGATTCACCTGTTCCAGAACGTGATTCCTCATCTACTGCGCT 700
|||||
234 erLeuSerGluTyrAspGlnCysLeuGluGluAsnGlnGluAsnArg 250
|||||
701 CCCTGAGCGAGTATGACCTGCTAGAGGAGATCGATCAGAGAACCGC 750
|||||
251 MetLysGluSerLeuAlaLeuPheGlyThrIleLeuGluLeuProTr 267
|||||
751 ATGGAGGAGAGTCTGCTGTTTCAGCACCATCTAGAGCTGCCCTGGTT 800
|||||
267 eLysSerThrValIleLeuPheLeuAsnLysThrAspIleLeuGlu 284
|||||
801 CAAGAGCAGCTCGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 850
|||||
284 LuLysIleProThrSerHisLeuAlaThrTyrPheProSerPheGln 300
|||||
851 ATAAGATTTCACCTCCACCTCCGCGCTGCGCACATCTTCTCCAGGGA 900

```

[illegible]

seq_name: qb_ro1:MUSGNA15A

seq_documentation_block:					
LOCUS	MUSGNAL5A	1353 bp	mRNA	ROD	27-APR-1993
DEFINITION	Mouse G protein alpha subunit (GNA-15) mRNA, complete cds.				
ACCESSION	M80632				
VERSION	M80632.1	GI:193570			
KEYWORDS	G protein alpha-subunit.				
SOURCE	Mus musculus (strain C57BL/6J x DBA F1) (library: lambda ZAP II) adult spleen cDNA to mRNA.				
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 1353)				
AUTHORS	Wilmon,T.M., Scherlie,P.A., Strathmann,M.P., Slepak,V.Z. and Simon,M.I.				
TITLE	Characterization of G-protein alpha subunits in the Gq class: Expression in murine tissues and in stromal and hematopoietic cell lines				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 88, 10049-10053 (1991)				

FEATURES

gene
CDS

[illegible]

```
alignment_scores:
  Quality: 1705.00
  Ratio: 4.658
  Percent Similarity: 97.861
```

```

alignment_block:
  US-08-878-801-2 x MUSGNA15A  ..
  Align seg 1/1 to: MUSGNA15A from: 1 to: 1353

1  MetAlaArgSerLeuThrTrpArgCysCysProTrpCysLeuThrGluAs 17
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
43  ATGGCCCGGTCCCTGACTGGTGGGCTGCTGCTCCCTGGCTGCACAGGA 92
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
17  pGluysAlaAlaAlaArgValAspGlnGluIleAsnArgIleLeuLeu 34
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
93  GGAAAGACTGCCGCCAGAATCGACAGGAGATCAACAGGATTTGTGTG 142
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
34  LuGlnLysGlnAspArgGlyGluLeuLysLeuLeuLeuGlyPro 50
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
143  AACAGAAAACAGAGCGCGAGAATTGAACCTCTCTGTTGGGGCCT 192
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
51  GlyGluSerGlyLysSerThrPheIleLysGlnMetArgIleIleHisG 67
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
193  GGTGAGAGCGGAAGATGACTTCATCAGCAGATGGCATCATTCACGG 242
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
67  yAlaGlyTyrSerGluGluGlyGlyGlyPheArgProLeuValTyrG 84
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
243  TGTGGGTACTCGAGAGGAGCCGACAGAGCTTCGCGCTGCTCATCTACC 292
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
84  InAsnIlePheValSerMetArgAlaMetIleGluAlaMetGluArgLeu 100
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
293  AGAATCATCTCGTCTCATGCGAGCCATGATAGATCGATGGACCGGCTG 342
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
101  GlnIleProPheSerArgProGluSerLysHisAlaSerLeuValMe 117
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
343  CAGATCCCTTCAGCAGGCTGCACAGCAGCAGCAGCCAGCCTAGTGAT 392
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
117  tSerGlnAspProTyrLysValThrPheGluLysArgTyrAlaAla 134
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
393  GACCCAGGACCCCTATAAAGTGAGCACATTCGAGAGCCATATGACAGT 442
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
134  laMetGlnTrpLeuTrpArgAspAlaGlyIleArgAlaCysTyrGluArg 150
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
443  CCATCGAGTACCTGTGGCGGAGCGGGCATCCGTGCTATGCTACGAGCGA 492
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
151  ArgArgGluPheHisLeuAspSerAlaValTyrTyrLeuSerHisLe 167
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
493  AGCGGTGAATTCACCTCTGGACTCCGCGGTGTATTACCTGTACACCT 542
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
167  uGluArgIleThrGluGluGlyTyrValProThrAlaGlnAspValLeu 184
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
543  GGAGCGCATATCAGAGACAGCTACATCCCACTGCGCAAGACGTGTGTC 592
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
184  rGSerArgMetProThrThrGlyLeaSnGluTyrCysPheSerValGln 200
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
593  GCAGTCGCATGCCACACAGGAGCATCAATGAGTACTGCTCTCCGTGAAG 642
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
201  LysThrAsnLeuArgIleValAspValGlyGlnLysSerGluArgLy 217
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
643  AAAACCAACTGCGCATCGTGATGTGGTGGCCAGAGGTACAGACGCTAG 692
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
217  sLysTrpIleHisCysPheGluAsnValIleAlaLeuIleTyrLeuAla 234
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
693  GAAATGGATTCACTGTTTCGAGAACGTGATTGCCCTCATCTACCTGGCCT 742
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
234  erLeuSerGluTyrAspGlnCysLeuGluAsnAsnGlnGluAsnArg 250
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
743  CCTGAGCGAGTATGACCATGTCCTAGAGGAGAACGATCAGAGAACCCG 792
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
251  MetLysGluSerLeuAlaLeuPheGlyThrIleLeuGluLeuProTrpPh 267
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
793  ATGAGGAGAGACTCGCTCTCTTCACGACGATCCTAGAGCTGCCCTGTT 842
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
267  eLysSerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuGlu 284
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
843  CAAGAGCACCTCGGTATCTCTTCCTCAACAGACGACATCTCTGGAG 892
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```


source

1. 1150
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/tissue_type="oocyte"
39. 1118
/codon_start=1
/product="guanine nucleotide regulatory protein Gq alpha subunit"
/protein_id="AA52189.1"
/db_xref="GI:505690"
/translation="MTLSIMACCLSEAEKARRINDEIEROLRRDKRDARRELKLL
LGTSGSGKSTFIKMRHIGSGYSDERKFTKLYQNIQFTAMQAMIRAMETLKTPYK
YEHNGKHALVREVDKVFASFNPDVDAIKYLNWDPGIEQECIDRRRYQLSDSKYY
LNDVRIATQGLYPTQDDVLRVETPTGIIIEYFDLSQVIFRMVDVGQSRERRKWIH
CPENVTSMFLVALSEYDOVLVESNENMEESKALFRTIITYPFQNSVILFLNKK
DLLEEKIMYSHLVDFPEYDGPQDAQAAREFILKMFVDLNPDSDKIITYSHFTCATDT
ENIRFVAVKDTILQLNLKEYNLV"
BASE COUNT 351 a 254 c 267 g 278 t
ORIGIN

CDS

alignment_scores:
Quality: 1057.00 Length: 362
Ratio: 3.571 Gaps: 2
Percent Similarity: 81.768 Percent Identity: 56.354
alignment_block:
US-08-878-801-2 x XLU10502 ..
Align seg 1/1 to: XLU10502 from: 1 to: 1150

13 CysLeuThrGluAspGluLysAlaAlaArgValAspGlnGluLeuAs 29
66 TGCCTCAGCGAGGAGGCCAAGGAAGCTCGAAGATTAAACGATGAGATCGA 115
29 nargileLeuLeuGluGlnLysGlnAspArgGlyGluLeuLysLeuL 46
116 GAGGAGCTCCCGCGGGGATAGAGAGAGCCCGCGGAGCTCAAACTCC 165
46 euLeuLeuGlyProGlyGluSerGlyLysSerThrPheileLysGlnMet 62
166 TCTGTCTGCGCACAGGAGAAAGTGGAAAAAGACACATTTATCAACAGATG 215
63 ArgileileHisGlyAlaGlyTyrSerGluGluArgLysGlyPheAr 79
216 AGGATAATCCATGGATCTGGTATTTCATGATGAGATAGAGGGGGTTCAC 265
79 gProLeuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluA 96
266 AAAGCTTGCTATCAAAATATATTCACATGCAATGCAAGCTATGATTCGGG 315
96 laMetGluArgLeuGlnIleProPheSerArgProGluSerLysHisHis 112
316 CCATGGAACCCCTTAAATACCATACATGATGATGAGATGCAACCAAGGCCAT 365
113 AlaSerLeuValMetSerGlnAspProTyrLysValThrThrPheGluLy 129
366 GCATCTGTAGTTCGAGAGTTGATGTAGAGAGGTGCGCTCTTTTGAGAA 415
129 sArgTyrAlaAlaAlaMetGlnTrpLeuTrpArgAspAlaGlyIleArgA 146
416 TCCTTATGTAGATGCAATAAAGTATTGTGGATGACCCCTGGCATACAAG 465
146 laCysTyrGluArgArgGlyPheileHisLeuLeuAspSerAlaValTyr 162
466 AGTGCTACGACGAGCGGAGGATATCATGCTGCTGATCCACTAATATAT 515
163 TyrLeuSerHisLeuGluArgIleThrGluGlyTyrValProThrAl 179
516 TATCTCAACGACGTGACCGCTATAGCAACTCAAGGATACCTTCCACCCA 565
179 aGlnAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTyrC 196

201 LysThrAsnLeuArgIleValAspValGlyGlyGlnLysSerGluArgLy 217
872 AAACCAAACTCGCATCGTGTGATGTTGGTGGCCAGAGGTTCAGACGATG 921
217 LysThrIleHisCysPheGluAsnValIleAlaLeuIleTyrLeuAlaS 234
922 GAAATGATTCATCTGTTTCGAGAAGCTGATTCCTCATCTACCTGGCT 971
234 erLeuSerGluTyrAspGlnCysLeuGluGluAsnGlnGluAsnArg 250
972 CCTGAGCAGATGATGACCATGCTGCTAGAGGAGAACATCAGAGAACCC 1021
251 MetLysGluSerLeuAlaLeuPheGlyThrIleLeuGluLeuProThrPh 267
1022 ATGAGGAGAGTCTCGCTCTGTTTCAGCAGCATCTAGAGCTGCCCTGGT 1071
267 eLysSerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuGluG 284
1072 CAAGACACCTCGGTCTCTCTCTCAACAGACGACATCCTGGAAG 1121
284 luLysIleProThrSerHisLeuAlaThrTyrPheProSerPheGlnGly 300
1122 ATAGATTACACCTCCACCTGCCACATACCTCCACAGTTCACAGGGA 1171
301 ProLysGlnAspAlaGluAlaLysArgPheIleLeuAspMetTyrTh 317
1172 CCCGCGGAGACGACAGAGCGCGCAAGAGCTTCATCTGACATGTATGC 1221
317 rArgMetTyrThrGlyCysValAspGlyProGluGlySerLysLysGlyA 334
1222 GCGGCTGACGCGAGCTGCCAGAGCCAGAGCCAGCGGTGGGAGGAGGCT 1271
334 laArgSerArgLeuLeuPheSerHisTyrThrCysAlaThrAspThrGln 350
1272 CCCGCGGCGCGCTTCTGCGACACTTCACCTGTGCCAGCGACAGCAA 1321
351 AsnIleArgLysValPheLysAspValArgAspSerValLeuAlaArgTy 367
1322 AGCGTCCGACGCGTTCAGGAGCTCGGAGCTCGGAGTCTGCTGCCCGGTA 1371
367 rLeuAspGluIleAsnLeuLeu 374
1372 CCTGGACGAGATCAACCTGCTG 1393
seq_name: gb_ov:XLU10502
seq_documentation_block:
LOCUS XLU10502 1150 bp mRNA VRT 04-NOV-1994
DEFINITION Xenopus laevis guanine nucleotide regulatory protein Gq alpha subunit mRNA, complete cds.
ACCESSION U10502.1 GI:505689
VERSION U10502.1 GI:505689
KEYWORDS
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 1150)
Shapira,H., Way,J., Lipinsky,D., Oron,Y. and Battey,J.F.
Neuromedin B receptor, expressed in Xenopus laevis oocytes, selectively couples to G alpha q and not G alpha 11 [published erratum appears in FEBS Lett 1994 Aug 1;349(2):318]
FEBS Lett. 348 (1), 89-92 (1994)
JOURNAL 94298961
MEDLINE
REFERENCE 2 (bases 1 to 1150)
Battey,J.F.
AUTHORS
TITLE Direct Submission
SUBMITTED (10-JUN-1994) James F. Battey, National Cancer Institute, Laboratory of Biological Chemistry, Bldg. 37, Room 5D-02, Bethesda, MD 20892, USA
JOURNAL
FEATURES Location/Qualifiers

```

566 GCAAGATGTGCTGAGAGTTCGAGTCCCAACACAGAGGAATCATTCAGATC 615
196 yspheSerValGlnLysThrAsnLeuArgLeuValaspValGlyGln 212
    |||.....|||.....|||.....|||.....|||.....|||.....|||
616 CCTTCGACTTACAAGTGTATCTTCAGAAATGGTGTAGTGTAGGAGGACAG 665
213 LysSerGluArgLysLysTrpIleHisCysPheGluAsnValIleAlaLe 229
    :|||.....|||.....|||.....|||.....|||.....|||.....|||
666 AGATCTGAGAGAGGAATGGATACACTGTTTGAACACGTCAACCTCAT 715
229 uileTyrLeuAlaSerLeuSerGluTyrAspGlnCysLeuGluGluAsnA 246
    :|||.....|||.....|||.....|||.....|||.....|||.....|||
716 AATGTTCTTGTAGTCCCTTAGTGAATATGATCAAGTCTGCTGCAATCTG 765
246 snGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThrIleLeu 262
    :|||.....|||.....|||.....|||.....|||.....|||.....|||
766 ACAATGAGAACCGAATGGAAGAAAGCAAGCGTTATTTAGGACGATCATC 815
263 GluLeuProTrpPheLysSerThrSerValIleLeuPheLeuAsnLysTh 279
    |||.....|||.....|||.....|||.....|||.....|||.....|||
816 ACATATCCGCTGTTCAAAATTCCTCCGTCATCTCTTCTTAACAGAA 865
279 rAspIleLeuGluGluLysIleProThrSerHisLeuAlaThrTyrPheP 296
    |||.....|||.....|||.....|||.....|||.....|||.....|||
866 AGATCTTCTAGAGGAAATCAATCATGCTACTCTGCTAGTGTATTTCC 915
296 roSerPheGlnGlyProLysGlnAspAlaGluAlaAlaLysArgPheIle 312
    |||.....|||.....|||.....|||.....|||.....|||.....|||
916 CTGATATGATGCGCCCAACCTGATGCAACAGCAGCAAGAAATTCATC 965
313 LeuAspMetTyrThrArgMetTyrThrGlyCysValaspGlyProGluG 329
    |||.....|||.....|||.....|||.....|||.....|||.....|||
966 TGAATATGTTGTAGATCTC.....ATAATCTACTCGCCTCACCTGCG 997
329 yserLysLysGlyAlaArgSerArgLeuPheSerHisTyrThrCysA 346
    :|||.....|||.....|||.....|||.....|||.....|||.....|||
998 CGACAAA.....ATAATCTACTCGCCTCACCTGCG 1029
346 laThrAspThrGlnAsnIleArgLysValPheLysAspValArgaspSer 362
    |||.....|||.....|||.....|||.....|||.....|||.....|||
1030 CCACAGATACTGAGAACATTCGTTCTGCTTGTGTCAGTCAAGGACACA 1079
363 ValLeuAlaArgTyrLeuAspGluIleAsnLeuLeu 374
    :|||.....|||.....|||.....|||.....|||.....|||.....|||
1080 ATCTACAGCTAAACCTGAGGAGTACACCTGGTG 1115

```

seq_name: gb_ov: XELGTPBIND

```

seq_documentation_block:
LOCUS      XELGTPBIND      1347 bp      mRNA      VRT      28-APR-1993
DEFINITION Xenopus laevis alpha subunit of Gq Gtp-binding protein mRNA,
            complete cds.
ACCESSION  L05540
VERSION    L05540.1 GI:214187
KEYWORDS   GTP-binding protein.
SOURCE     Xenopus laevis (library: lambda-gtl0 from Doug Melton) ovary cDNA
            to mRNA.
ORGANISM   Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
            Xenopus.
REFERENCE  1 (bases 1 to 1347)
AUTHORS   Lader, K.J. and Smith, L.D.
TITLE     Isolation of the Xenopus cDNA for G alpha q
JOURNAL   Unpublished (1992)
FEATURES   Location/Qualifiers
            1..1347
            /organism="Xenopus laevis"
            /db_xref="taxon:8355"
            /tissue_type="ovary"
            /tissue_lib="lambda-gtl0 from Doug Melton"
            1..1080
            /codon_start=1
CDS

```

```

/product="GTP-binding protein"
/db_xref="AAA49730.1"
/db_xref="GI:214188"
/translation="MTLESIWACCLSEAEARRINDEIRDLRDKRDARRELKLL
LGTSGSKSTFKOMRIIHGSCYSDKRGFTKLVYQNTFSAMQAMIRAMETLKI
YK YEHNGHALLVREVDVEXVAFENPYDAIKYLMNDPGIQECYDRREYQ
LNDLDRATHGYLPQDVLVRVPTTGIIETFPDQSVIFRMVDVGORSRRKWI
CFENVTSMFLVALSEYDQVLVESDNENRMEESKALFRTIITYPMFQNSV
DLLEKINYSHLVDYFPYDQPDQAAREFIILKMFVDLPDSDSKIIYSHT
ENIRFVFAAVKDTILQLNLKSYNLV"
<136..>156
/bound_moiety="GTP"
1081..>1347
protein_bind      415 a      282 c      306 g      344 t
3'UTR
BASE COUNT
ORIGIN
alignment_scores:
    Quality: 1052.00      Length: 362
    Ratio: 3.566      Gaps: 2
    Percent Similarity: 81.492      Percent Identity: 56.354
alignment_block:
US-08-878-801-2 x XELGTPBIND
Align seg 1/1 to: XELGTPBIND from: 1 to: 1347

```

```

13 CysLeuThrGluAspGluLysAlaAlaAlaArValaspGlnGluIleAs 29
    |||.....|||.....|||.....|||.....|||.....|||.....|||
28 TGCTTCAGGAGGAGCGGAGGAGCTCGAAGCATTAACGATGAGATCGA 77
    |||.....|||.....|||.....|||.....|||.....|||.....|||
29 nArgIleLeuLeuGluGlnLysLysGlnAspAlaGlyGluLeuLysLeu 46
    |||.....|||.....|||.....|||.....|||.....|||.....|||
78 GAGGAGCTCCCGCGGATAAGAGAGAGCCGCCCGGGAGCTCAAACTCC 127
    |||.....|||.....|||.....|||.....|||.....|||.....|||
46 euLeuLeuGlyProGlyGluSerGlyLysSerThrPheIleLysGlnMet 62
    |||.....|||.....|||.....|||.....|||.....|||.....|||
128 TCCTCTCGGCACAGGAGAAAGTGGAAAAGGCACATTATCAACAGATG 177
    |||.....|||.....|||.....|||.....|||.....|||.....|||
63 ArgIleIleHisGlyAlaGlyTyrSerGluGluArgLysGlyPheAr 79
    |||.....|||.....|||.....|||.....|||.....|||.....|||
178 AGGATAATCCATGGATCTGGTTATTTCAGATGAAGATAAGAGGGGTTCAC 227
    |||.....|||.....|||.....|||.....|||.....|||.....|||
79 gProLeuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluA 96
    |||.....|||.....|||.....|||.....|||.....|||.....|||
228 AAAGCTTGTCTATCAAAATATATTCTCTGCAATGCAAGCTATGATTCGG 277
    |||.....|||.....|||.....|||.....|||.....|||.....|||
96 laMetGluArgLeuGlnIleProPheSerArgProGluSerLysHisHis 112
    |||.....|||.....|||.....|||.....|||.....|||.....|||
278 CCATGGAACCCCTTAAATACCATACAGATGATGAACACACAAAGGCCAT 327
    |||.....|||.....|||.....|||.....|||.....|||.....|||
113 AlaSerLeuValMetSerGlnAspProTyrLysValThrPheGluLys 129
    |||.....|||.....|||.....|||.....|||.....|||.....|||
328 GCACCTGTTAGTTCGAGAAGTTGATGTAGAGAAGATCGCCTCTTTTGAGAA 377
    |||.....|||.....|||.....|||.....|||.....|||.....|||
129 sArgTyrAlaAlaAlaMetGlnTrpLeuTrpArgAspAlaGlyIleArg 146
    |||.....|||.....|||.....|||.....|||.....|||.....|||
378 TCCTTATGATGTCATAAAGTATTGTGGAAATGACCCCTGGCATACAG 427
    |||.....|||.....|||.....|||.....|||.....|||.....|||
146 laCysTyrGluArgArgGluPheHisLeuLysAspSerAlaValTyr 162
    |||.....|||.....|||.....|||.....|||.....|||.....|||
428 AGTCTACGACAGCGGACGGATATCATGCTGTGTGATCCACTAAATAT 477
    |||.....|||.....|||.....|||.....|||.....|||.....|||
163 TyrLeuSerHisLeuGluArgIleThrGluGluLutlyTyrValProThrAl 179
    |||.....|||.....|||.....|||.....|||.....|||.....|||
478 TATCTCAATGACCTGGACCTATAGCAACTCATGCTACCTCTCTACCCA 527
    |||.....|||.....|||.....|||.....|||.....|||.....|||
179 aGlnAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTyrC 196
    |||.....|||.....|||.....|||.....|||.....|||.....|||
528 GCAAGATGCTGAGAGTTCGAGTCCCAACACAGCGGAATCATTCAGTATC 577
    |||.....|||.....|||.....|||.....|||.....|||.....|||
196 yspheSerValGlnLysThrAsnLeuArgIleValaspValGlyGln 212
    |||.....|||.....|||.....|||.....|||.....|||.....|||

```

```

578 CCTTTGACTTACAAAGCGTATCTTCAGATGGTGGATGTAGGAGACAG 627
213 LysSerGluArgLysIleHisCysPheGluAsnValIleAlaLe 229
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
628 CGCTCTGAGAGGAAGAAATGGATACACTGCTTGAATGTCACCTCAT 677
229 uIeTyLeuAlaSerLeuSerGluTyAspGlnCysLeuGluGluAsnA 246
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
678 AATGTTCTTGTAGCCCTTAGTGAATATGATCAAGTGTGGTGAATCTG 727
246 snGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThrIleLeu 262
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
728 ACAATGAGAACCGATGGAAGAAAGCAAGCGTTATTAGGAGCATCATC 777
263 GluLeuProTPrpPheLysSerThrValIleLeuPheLeuAsnLysTh 279
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
778 ACATATCCATCGTTTCAAAATCTCAGTCATCTGTTTCTAAACAAGAA 827
279 rAspIleLeuGluGluLysIleProThrSerHisLeuAlaThrTyPheP 296
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
828 AGATCTTCTAGAGAAATAATCATGCTACTGCTAGTGTGATTTATTC 877
296 roSerPheGlnGlyProLysGlnAspAlaGluAlaAlaLysArgPheIle 312
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
878 CTGAATATGATGACCCCGAGAGTGCACAGCAGCAAGAGAAATTCATC 927
313 LeuAspMetTyThrArgMetTyThrGlyCysValAspGlyProGluGl 329
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
928 TTAAGATGTTCTAGATCTC.....AATCCAGACAG 959
329 ySerLysLysGlyAlaArgSerArgArgLeuPheSerHisTyThrCysA 346
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
960 CGACAAA.....ATAATCTACTGCACCTTACCTGTG 991
346 laThrAspThrGlnAsnIleArgLysValPheLysAspValArgAspSer 362
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
992 CTACAGATACAGAACATCTGTTCTCTTGTGTCAGTCAAGGACACA 1041
363 valLeuAlaArgTyLeuAspGluIleAsnLeuLeu 374
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1042 ATCCTACAGCTAAATCTGAAGAGATACAACTGGTG 1077

```

seq_name: gb_om:DOGGALP

```

seq_documentation_block: 1077 bp mRNA MAM 02-AUG-1996
LOCUS DOGGALP
DEFINITION Canis familiaris G alpha q mRNA, complete cds.
ACCESSION L76257
VERSION L76257.1 GI:1478043
KEYWORDS G alpha q; coupling; thromboxane receptor.
SOURCE Canis familiaris
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 1077)
AUTHORS Dunlop,P.C., Leis,L.A. and Johnson,G.J.
TITLE Canine Thromboxane Receptor-G Protein Coupling
JOURNAL Unpublished (1996)
FEATURES
    source
        1..1077
            /organism="Canis familiaris"
            /db_xref="taxon:9615"
            /cell_type="platelet"
            /dev_stage="adult"
            <1..>1077
            /gene="G alpha q"
            /gene="G alpha q"
            /gene="G alpha q"
            /codon_start=1
            /protein_id="AAB05548.1"
            /db_xref="GI:1478044"
mRNA
gene
CDS

```

/translation="MTLESIMACCLSEAEARRINDEIEROLRRDKRAREKLLLL
LGTGSGSKSTIKOMRIHSGYSDERKFTKLIVYQNIPTAMQAMIRAMDTLKIPYK
YHNKHAQALYREVDEVEKSAFENPYDAIKLWNPQIOECYDRRREYQLSDSTKY
LELDORVADPAILPQQOYLVRVPTTGLIEVPDLOSIFRMVDVGQRRRKKWIH
CFENVTSIMFLVALSEYDOVLVESDNENRMEESKALFTIITYPWFQNSVLLFLNKK
DLLEEKIMYSHLVDFEYDPGPDQAQAAREFILKMFVDLNPDSKLIYSHTCATDT
ENIRFVFAAVKDTILQLNLKEYNLV"

BASE COUNT 311 a 259 c 266 g 241 t

ORIGIN

alignment_scores:
Quality: 1047.00 Length: 362
Ratio: 3.537 Gaps: 2
Percent Similarity: 81.768 Percent Identity: 55.801

alignment_block:

US-08-878-801-2 x DOGGALP ..

Align seg 1/1 to: DOGGALP from: 1 to: 1077

```

13 CysLeuThrGluAspGluLysAlaAlaAArgValAspGlnGluIleAs 29
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
28 TGCCTGACGAGGAGGCCAAGAGCCCGCGGATCAACGACGAGATCGA 77
29 nArgIleLeuLeuGluGlnLysLysGlnAspArgGlyGluLeuLysLeu 46
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
78 GCGGCAGCTCCGACGAGCAAGCGGCGCGCGGAGCTCAAGCTGC 127
46 euLeuLeuGlyProGlyGluSerGlyLysSerThrPheIleLysGlnMet 62
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
128 TGCTGCTCGGGACGGGAGAGAGTGGCAAGAGTACGTTTATCAAGCAGATG 177
63 ArgIleIleHisGlyAlaGlyTySerGluGluGluArgLysGlyPheAr 79
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
178 AGAATCATCCATGGTGGGCTACTCTGATGAAGACAAAAGGGGCTTCAC 227
79 gProLeuValTyGlnAsnIlePheValSerMetArgAlaMetIleGluA 96
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
228 CAAGTGTGTATCAGAACATCTTCAGGCCCATCGAGGCTATGATCAGAG 277
96 laMetGluArgLeuGlnIleProPheSerArgProGluSerLysHisHis 112
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
278 CCATGGACACGCTCAAGATCCCTACAGTACGAGCACAATAAGGCTCAT 327
113 AlaSerLeuValMetSerGlnAspProTyLysValThrThrPheGluLys 129
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
328 GCACAATTAGTTGAGAGTCGACGCTGGAGAGGTGCTGCTTTTGAGAA 377
129 sArgTyAlaAlaAlaMetGlnTriLeuTriArgAspAlaGlyIleArgA 146
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
378 TCCATATGTAGACCAATAAAGAGTTTATGGAATGATCTCGGAATCCAGG 427
146 laCysTyArgGluArgArgGluPheHisLeuLeuAspSerAlaValTy 162
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
428 AATGCTATGACACACGACGAGATATCAGTTGTCGCACTCTACCAATAC 477
163 TyLeuSerHisLeuGluArgIleThrGluGluGlyTyValProThrAl 179
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
478 TATCTTAATGACTTGGACCGCTGTGTCGCCCTTACCTGCTGCTAGCA 527
179 aGlnAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTy 196
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
528 ACAAGATGTCTTAGAGTTCGAGTCCCGCCACACAGGAGTATCGAATACC 577
196 ysPheSerValGlnLysThrAsnLeuArgIleValAspValGlyGln 212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
578 CCTTTGACTTACAAGTGTCTATTTTCAAGATGGTTCGATAGGGGGCCAG 627
213 LysSerGluArgLysLysTyPheIleHisCysPheGluAsnValIleAlaLe 229
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
628 AGGTCAGAGAGAGAAATGGATACACTGCTTTTGAAATGTACCTCTAT 677

```


BASE COUNT 309 a 255 c 276 g 240 t
ORIGIN
DLEKIMVSHLVDPYFPEYDGPQDQAAREFLKMFVDLNPDSKIIYSHFTCATDT
ENIRFVFAAVKDTILOLNKYNLV"

alignment_scores:
Quality: 1046.00 Length: 362
Ratio: 3.534 Gaps: 2
Percent Similarity: 81.768 Percent Identity: 55.525

alignment_block:
US-08-878-801-2 x AF2344260 ..
Align seg 1/1 to: AF2344260 from: 1 to: 1080
13 CysteLeuThrGluAspGluLysAlaAlaAargValAspGlnGluLeuAs 29
28 TGCCTGAGCGAGGAGCCAAAGAGCCCGGAGGATCAACGACGAGTCGA 77
29 nArgileLeuLeuGluGlnLysLysGlnAspArgGlyGluLeuLysLeu 46
78 CGCGCAGCTGCGCAGGACGACAGCCGCGGAGCTCAAGCTGC 127
46 euLeuLeuGlyProGlyGlySerGlyLysSerThrPheLeuLysGlnMet 62
128 TGCTGCTGGGACAGGAGGAGAGTGGCAAGAGTACCTTCATTAAAGCAGAT 177
63 ArgileLeuHisGlyAlaGlyTyrSerGluGluGluArgLysGlyPheAr 79
178 AGGATCATCCACGCGGTGGGTACTCTGTGATGACAGACAGAGGGGCTTAC 227
79 gProleuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluA 96
228 CAAACTGGTGTATCAGAACATCTTTACAGCCATGCAGGCCATGTGCAGAG 277
96 laMetGluArgLeuGlnIleProPheSerArgProGluSerLysHisHis 112
278 CTATGGACACTCTCAAGATCCCATACAAAGTATGACACCAATAAGGCTCAT 327
113 AlaSerLeuValMetSerGlnAspProTyrLysValThrPheGluLys 129
328 GCACAAATGGTTCGAGAGGTTGATGTGGAGAAGGTCTCTGTTTGGAGAA 377
129 sArgTyrAlaAlaAlaMetGlnTrpLeuTrpArgAspAlaGlyIleArgA 146
378 TCATATATGATGAGCAATTAAGAGCTTGTGAATGATCTCGAATCCAGG 427
146 laCystTyrGluArgArgGluPheHisLeuLeuAspSerAlaValTyr 162
428 AATGCTACATAGACGGCGAGAAATATCAGCTATCTGACTCTACCAATAC 477
163 TyrLeuSerHisLeuGluArgIleThrGluGluGlyTyrValProThrAl 179
478 TATCTGAACGACTTGGACCGTGTGGCTGACCTCTCTATCTGCTACACAA 527
179 aGlnAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTyrC 196
528 ACAAGATGTCTTAGAGTTCGAGTCCCGCCACAGGGATCATTTAGTACC 577
196 yPheSerValGlnLysThrAsnLeuArgIleValAspValGlyGlyGln 212
578 CCTTCGACTTACAGAGTGTCACTCTTCAGAAATGCTGATGTAGGAGCCAA 627
213 LysSerGluArgLysLysTrpIleHisCysPheGluAsnValIleAlaLe 229
628 AGGTCAGAGAGAGAAATGGATACACTTGTCTTTGAAACGCTCACCTCGAT 677
229 uIleTyrLeuAlaSerLeuSerGluTyrAspGlnCysLeuGluGluAsnA 246
678 CATGTTCTGTCGCGCTTAGCGAATACGATCAATGTTCTTGTGGAGTCAG 727
246 snGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThrIleLeu 262

246 snGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThrIleLeu 262
.....
728 ACATGACAGACCGATGGAGGAGGAGGCTCTTTAGAACAAATTATC 777
263 GluLeuProTyrPheLysSerThrSerValIleLeuPheLeuAsnLysTh 279
.....
778 ACATACCCCTGTTCCAGAACTCCCTGGTTATCTGTCTTTAAACAAGAA 827
279 rAspIleLeuGluGluLysIleProThrSerHisLeuAlaThrTyrPheP 296
.....
828 AGATCTTCTAGAGGAGAAATCATCTATCCCACTAGTCGACTACTTCC 877
296 roSerPheGlnGlyProLysGlnAspAlaGluAlaAlaLysArgPheIle 312
.....
878 CAGATATGATGACCCAGAGAGATGCCAGGACGAGCCGAGAAATTCAT 927
313 LeuAspMetTyrArgMetTyrThrGlyCysValAspGlyProGluGlu 329
.....
928 CTCAAGATGTCGTGGAGCTG.....AACCCAGACAG 959
329 ySerLysLysGlyAlaArgSerArgArgLeuPheSerHisTyrThrCysA 346
.....
960 TGACAAA.....ATTATCTACCTCCCACTTCACGTGCG 991
346 laThrAspThrGlnAsnIleArgLysValPheLysAspValArgAspSer 362
.....
992 CCACAGACACCGAGATATCCGCTTTGTCTTGTGCTGCCGTCAAGGACAC 1041
363 valLeuAlaArgTyrLeuAspGluIleAsnLeuLeu 374
.....
1042 ATCTCCAGTTGAACCTGAAGGAGTACATCTGTGTC 1077

seq_name: gb_rol:AF234260

seq_documentation_block: 1080 bp mRNA ROD 26-MAR-2000
LOCUS AF234260
DEFINITION Rattus norvegicus heterotrimeric guanine nucleotide-binding protein
alpha q subunit mRNA, complete cds.

ACCESSION AF234260
VERSION AF234260.1 GI:7329186
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Rattus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE 1 (bases 1 to 1080)
AUTHORS Strothmann, R.
TITLE Rat G alpha q subunit
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1080)
AUTHORS Strothmann, R.
TITLE Direct Submission
JOURNAL Submitted (14-FEB-2000) Pharmacology, Freie Universitaet Berlin,
Thielallee 69-73, Berlin 14195, Germany

FEATURES
source
1..1080
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/tissue_type="brain"
1..1080
/note="G protein"

CDS
1..1080
/codon_start=1
/product="heterotrimeric guanine nucleotide-binding
protein alpha q subunit"
/protein_id="AAF59930.1"
/db_xref="GI:7329187"
/translation="MTLESIACCLSEAEKRRINDEITERLQRRDKRDARRELKLL
LGESGKSTFFKQMRILHSGVSDKRGFTKLVYONIFTAMQAVRAMDTLKIPYK
YEHKAGQLREVDVEKVSAPENPYDAIKSLWNPDPGIECYDRREYQLSDSTKYI
LNDLRVADPSYLPDQVLRVPTGPIETFPDQLQSIVFRWVDVGORSERRKWIH
CFENVTISIMFValseyDQVLVESDNENRMEESKALFRLTIITYPWFQNSVILFNKK

```

728 ACAATGAGAACCAAGTGGAGGAGCAAGCACTCTTTAGAACCATATC 777
263 GluLeuProThrPheLysThrSerValIleLeuPheLeuAsnLysTh 279
778 ACATATCCCTGGTTCAGAACCTCTCTGTATTCTCTTAAACAAGAA 827
279 rAspIleLeuGluGluLysIleProThrSerHisLeuAlaThrTyrPhe 296
828 ACATCTTCAGAGGAGAAATATGATTCCACCTAGTCGACTACTTCC 877
296 roSerPheGlnGlyProLysGlnAspAlaGluAlaLysArgPheIle 312
878 CAGATATATGAGCCAGAGAGATGCCAGGACGACGAGCAATTCATC 927
313 LeuAspMetTyrThrArgMetTyrThrGlyCysValAspGlyProGlu 329
928 CTGAAGATGTCCTGGACCTG.....AACCCGACAG 959
329 ySerLysLysGlyAlaArgSerArgArgLysPheSerHisTyrThrCys 346
960 TGACAAA.....ATCATCTACTCGCACTTCAGCTGTG 991
346 laThrAspThrGlnAsnIleArgLysValPheLysAspValArgSpSer 362
992 CCACAGACAGAGAACATCCGCTCGTGTTCCTGCTGTCAGGACACC 1041
363 ValLeuAlaArgTyrLeuAspGluIleAsnLeuLeu 374
1042 ATCTGTCAGCTGAACCTGAGGAGTACATCTGGTC.1077

```

seq_name: gb_om:BOVGL2A

```

seq_documentation_block: 1391 bp mRNA MAM 07-FEB-1999
LOCUS BOVGL2A
DEFINITION Bovine mRNA for GTP-binding protein alpha-subunit (GL2A), complete
ACCESSION D90336
VERSION 1
KEYWORDS G protein; GL2-alpha; GTP-binding protein; GTPase.
SOURCE Bovine liver, cDNA to mRNA, clone pGL7.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
Nakamura, F., Ogata, K., Shiozaki, K., Kameyama, K., Ohara, K., Haga, T.
and Nakada, T.
Identification of two novel GTP-binding protein alpha-subunits that
lack apparent ADP-ribosylation sites for pertussis toxin
J. Biol. Chem. 266 (19), 12676-12681 (1991)
91286303
These data kindly submitted in computer readable form by: Fumio
Nakamura
Department of Biochemistry
Institute of Brain Research
Faculty of Medicine, University of Tokyo
7-3-1 Hongo, Bunkyo-ku
Tokyo 113
Japan
phone: 03-3812-2111 x3561
Fax: 03-3814-8154.
Location/Qualifiers
1..1391
/organism="Bos taurus"
/db_xref="taxon:9913"
251..1312
/codon_start=1
/feature="GTP-binding protein, alpha-subunit"
/protein_id="BAA14350.1"
/db_xref="GI:217568"
/translation="MACCLSDVYKSKRIAEIKQLRRDKRDARRELKLLLLTGTGES
GKSTFIKQMRILHGAGYSEEDKRGFTKLIVQNIPTAQAMIRAMETKLILKYEQNKA

```

FEATURES

source

CDS

```

NALLIREVDVKKVTEEHRYVSAIKTLWDPGICQCYDRREYQLSDSAKYYLTVDVR
IATSGYLPQODVLVRVPTGLIEYFDLENIIFRMVDVGGQSRERRKWIHCFENVT
SIMPLVASEYDQVLVESDNIINRMEESKALERTIVTFWPNSSVILFNKRDLEDK
ILHSHLVDFEPFEGRPORDAKAAREFILKMFVDLNPDSKLIYSHTCATDTENIRFY
FAVKDRITLQNLNLEYNLV" 212 t
BASE COUNT 282 a 426 c 471 g
ORIGIN

```

alignment_scores:

Quality: 1046.00 Length: 362
Ratio: 3.570 Gaps: 2
Percent Similarity: 80.939 Percent Identity: 56.077

alignment_block:

US-08-878-801-2 x BOVGL2A ..

Align seg 1/1 to: BOVGL2A from: 1 to: 1391

```

13 CysLeuThrGluAspGluLysAlaAlaArgValAspGlnGluIleAs 29
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
260 TGCCTGAGCGATGAGGTGAAGGAGTCCAAGCGGATCAACGCCGAGATCGA 309
29 nArgIleLeuLeuGluGlnLysLysGlnAspArgGlyGluLeuLysLeu 46
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
310 GAAGCAGCTGCGCGGCGGACAAAGCGGACGCCGCGGAGCTCAAGCTGC 359
46 euleLeuGlyProGlyGluSerGlyLysSerThrPheIleLysGlnMet 62
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
360 TGCTGCTCGGACGCGGAGAGCGGGAAGAGCAAGTTCATCAAGCAGATG 409
63 ArgIleIleHisGlyValArgLysSerGluGluGluArgLysGlyPheAr 79
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
410 CGCATCATCCAGCGGCGGCTACTCAGAGAGGACAAAGCGGCGGCTTCAC 459
79 gProLeuValTyrGlnAsnIlePheValSerMe:ArgAlaMetIleGluA 96
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
460 CAAACTGCTGTACCAGAACATCTTACCCGCCATCAGGCCCATGATCCGCG 509
96 laMetGluArgLysGlnIleProPheSerArgP:ogLysSerLysHisHis 112
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
510 CCATGGAAACCTGAAGATCTCTACAAAGTACGACAGCAAGAACGCCCAAC 559
113 AlaSerLeuValMetSerGlnAspProTyrLysValThrPheGluLys 129
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
560 GCATCTCTGATCCGGGAGGTGGATGTGGAAAAGCTGACGACCTTCGAGCA 609
129 sArgTyrAlaAlaAlaMetGlnTrpLeuTrpArgAspAlaGlyIleArg 146
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
610 CCGGTACGTGAGCGCCCATCAAGACCTCTGTGCAACGACCCCGCATCCAGG 659
146 laCystTyrGluArgArgGluPheHisLeuLeuAspSerAlaValTyr 162
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
660 AGTGCTATGACCGCGCGGAGTACCAGCTCTCGACTCCGCCCAAGTAC 709
163 TyrLeuSerHisLeuGluArgIleThrGluGluLysTyrValProThrAl 179
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
710 TACCTGACGCGACGTGGACCGCATTCGCCACCTCAGCTACCTGCCACCCA 759
179 aGlnAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTyrC 196
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
760 GCAGGACGTGCTCGCGTGCCTGCCACCGGATCATCGAATACC 809
196 yPheSerValGlnLysThrAsnLeuArgIleValAspValGlyGlyGln 212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
810 CCTTCACCTGGAGAACATCATCTTCAGGATGGTGTGATGTGGGGGCCAG 859
213 LysSerGluArgLysLysTrpIleHisCysPheGluAsnValIleAlaLe 229
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
860 AGGTCCGCGGAGGAGGATGATTCATCTGCTTTGAGAACGTCAGCTCCAT 909
229 ulLeuTyrLeuAlaSerLeuSerGluTyrAspGlnCysLeuGluGluAsnA 246
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Thu Sep 6 15:16:47 2001

```

910 CATGTCCTTGTGGCCCTAGTACGACCAAGTGTGGTGAATCGG 959
246 snGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThrIleLeu 262
960 ACAACGAGAACCGCATGGAGGAGCAAGCGCTGTCCGGACCATCGTC 1009
263 GluLeuProThrPheLysSerThrSerValIleLeuPheLeuAsnLysTh 279
1010 ACTTACCCCTGGTTCAGAACTCATCCGTCATCTCTCTCAACAAGAA 1059
279 rAspIleLeuGluGluLysIleProThrSerHisLeuAlaThrTyrPheP 296
1060 GGACCTGCTGGAGGACAAAGTCTCCACTCCACCTGGTGACTACTTCC 1109
296 roSerPheGlnGlyProGluAspAlaGluAlaAlaLysArgPheIle 312
1110 CGGAGTTCGAGGCGCCCGAGGAGCGGAGCTGCCCGGAGTTTCATC 1159
313 LeuAspMetTyrThrArgMetTyrThrGlyCysValAspGlyProGluG 329
1160 CTGAAGATGTTCTGGGACCTA.....AACCCGACAG 1191
329 ySerLysLysGlyAlaArgSerArgLeuPheSerHisTyrThrCysA 346
1192 CGACAAG.....ATCATCTACTCGACCTTCACCTCGG 1223
346 laThrAspThrGlnAsnIleArgLysValPheLysAspValArgAspSer 362
1224 CCACCGACACAGACAGACATCGTTCGCTGCTGCTCAAGGACACCC 1273
363 ValLeuAlaArgTyrLeuAspGluIleAsnLeuLeu 374
1274 ATCTTCGACCTCAACCTGAAGGAGTACAACTGGTG 1309

seq_name: gb_pr10.HS043083

seq_documentation_block: 1700 bp mRNA PRI 30-AUG-1996
LOCUS HS043083
DEFINITION Human G alpha-q (Gaq) mRNA, complete cds.
ACCESSION U43083
VERSION U43083.1 GI:1174071
KEYWORDS
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1700)
Chen,B., Leverette,R.D., Schwinn,D.A. and Kwatra,M.M.
Human G(alpha q): cDNA and tissue distribution
Biochim. Biophys. Acta, Gene Struct. Expr. 1281 (2), 125-128 (1996)
96256639
2 (bases 1 to 1700)
Kwatra,M.M., Schwinn,D.A., Leverette,R.D. and Chen,B.
Direct Submission
Submitted (11-DEC-1995) Madan M. Kwatra, Anesthesiology, Duke
University Medical Center, Box 3094, Durham, NC 27710, USA
Location/Qualifiers
1..1700
/organism="Homo sapiens"
/db_xref="taxon:9606"
221..1300
/gene="Gaq"
221..1300
/gene="Gaq"
/codon_start=1
/product="G alpha-q"
/db_xref="GI:1174072"
/translation="MTLEISMACCLSEEAKERRINDETERLRDRKDRRELKILL
LGTGSKSTFKQMRIRIHSGSYDEKRGFTKLYFQNIPTAQAMIRAMDTLKIYK
YHNKAHOLRVEDYKVSFAFENPVYDAIKSLWNPDPGIECDRRRQLSDSKYI
LNDLRVADPAPLPTQDVLVRVPPTGTIIEYFDLQSVIRFWDVGGGRRRRKH
CFENVTSIMFLVSEYDQVLVESDNNRMEESKALFTIITYPWFQNSVILFLNKK

```

```

BASE COUNT 440 a 426 c 423 g 411 t
ORIGIN
alignment_scores:
Quality: 1045.00 Length: 362
Ratio: 3.530 Gaps: 2
Percent Similarity: 81.768 Percent Identity: 55.801
alignment_block:
US-08-878-801-2 x HS043083 ..
Align seg 1/1 to: HS043083 from: 1 to: 1700
13 CysLeuThrGluAspGluLysAlaAlaAlaArgValAspGlnGluLeu 29
248 TGCTGAGGAGGAGGCGCAAGGAGCCCGCGGATCAACGACGAGATCGA 297
29 nArgIleLeuLeuGluGlnLysLysGlnAspArgGlyGluLeuLysLeu 46
298 CGCGCAGCTCCGCGAGGACAAAGCGGAGCGCCGCGGAGCTCAAGCTGC 347
46 euLeuLeuGlyProGlyGluSerGlyLysSerThrPheIleLysGlnMet 62
348 TGCTGCTCGGCGACAGGAGAGTGGCAAGAGTACGTTTATCAAGCAGATG 397
63 ArgIleIleHisGlyAlaGlyTyrSerGluGluGluArgLysGlyPheAr 79
398 AGAATCATCCATGGTCAGGATCTCTGATGAAGATAAAGGGGCTTCAC 447
79 gProLeuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGlu 96
448 CAAGCTGCTGTATCAGAACATCTTCACGGCCATGTCAGGCCATGATCAG 497
96 laMetGluArgLeuGlnIleProPheSerArgProGluSerLysHisHis 112
498 CCATGGGACACTCAAGATCCATCAAGATGATGAGCACAAATAGGCTCAT 547
113 AlaSerLeuValMetSerGlnAspProTyrLysValThrThrPheGlyL 129
548 GCACAATTAGTTCGAGAAAGTTGATGTGGAGAAAGGTGCTGCTTTTGAG 597
129 sArgTyrAlaAlaAlaMetGlnTrpLeuTrpArgAspAlaGlyIleArg 146
598 TCCATATGATGATGCAATAAAGAGTTTATGAATGATCCTGGAATCCAG 647
146 laCysTyrGluArgArgGluPheHisLeuLeuAspSerAlaValTyr 162
648 AATGCTATGATAGACGACGAGAAATATCAATTATCTGACTCTACCAATAC 697
163 TyrLeuSerHisLeuGluArgIleThrGluGluGlyTyrValProThrAl 179
698 TATCTTAATGACTTGGACCGCGTAGCTGACCCCTGCTACCTGCCTACGCA 747
179 aGlnAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTyrC 196
748 ACAAGATGCTGTAGAGTTTCGAGTCCCGCCACACAGGAGTATCGAATACC 797
196 ysPheSerValGlnLysThrAsnLeuArgIleValAspValGlyGlyGln 212
798 CCTTTCAGTACAAAGTGTCAATTTTCAGAAATGTCACCTCTAT 847
213 LysSerGluArgLysLysTrpIleHisCysPheGluAsnValIleAlaLe 229
848 AGTCACAGAGAAAGAAATGATACACTGCTTTTGAATAATGTCACCTCTAT 897
229 uIleTyrLeuAlaSerLeuSerGluTyrAspGlnCysLeuGluGluAsnA 246
898 CATGTTTCTAGTACGCGCTTAGTGAATATGATCAAGTTCTGCTGGATCAG 947
246 snGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThrIleLeu 262

```

```

948 ACAATGAGAACCGAATGAGGAGAAAGCAAGGCTCTCTTTAGAACAAATTATC 997
263 GluLeuProTrpPheLysSerThrSerValIleLeuPheLeuAsnLysTh 279
998 ACATACCCCTGGTTCCAGAACTCCTCGTTATTCTGTCTTAAACAAGAA 1047
279 rAspIleLeuGluGluLysIleProThrSerHisLeuAlaThrTyrPheP 296
1048 AGATCTTCTAGAGAGAGAAATCATGTATTCCTCATCTAGTCGACTACTTC 1097
296 rOSerPheGlnGlyProLysGlnAspAlaGluAlaAlaLysArgPheIle 312
1098 CRGATATGATGAGCCCGACGAGAGATGCCCGAGCGCGGAGAAATTCATT 1147
313 LeuAspMetTyrThrArgMetTyrThrGlyCysValAspGlyProGluG 329
1148 CTGAAGATGTCTGTGACCTG.....AACCAGACAG 1179
329 ySerLysLysGlyAlaArgSerArgLeuPheSerHisTyrThrCysA 346
1180 TGACAAATTAAC.....TACTCCCACTTCAGGTGCG 1211
346 lathrAspThrGlnAsnIleArgLysValPheLysAspValArgAspSer 362
1212 CCACACACACCGAGAAATATCCGCTTTGCTTTGCTGCCGCTCAAGGACACC 1261
363 ValLeuAlaArgTyrLeuAspGluIleAsnLeuLeu 374
1262 ATCTCCAGTTGAACCTGAAGGAGTACAAATCTGGTC 1297
```


OM of: US-08-878-801-2 to: N_Geneseq_0601.* out_format : pfs
Date: Sep 6, 2001 1:04 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q/cgn2.1/USPTO.spool/6004808/runat_06092001_110135_4516/app_query.fasta_1.437
-DB=N_Geneseq_0601 -QFMT=fastap -SUFFIX=p2n.rng -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -XGAPEXT=0.500
-XGAPOP=4.500 -XGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=200000000
-USER=6004808 -CGWL_1_293 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT_THREADS=1

Search information block:
Query: US-08-878-801-2
Query length: 374
Database: N_Geneseq_0601.*
Database sequences: 730101
Database length: 313950809
Search time (sec): 77.720000

score_list:	Strd Orig	zScore	EScore	Len	Documentation
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAA48751	1360.00	3777.99	1.8e-202	2060	
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV03464	1347.00	3759.41	2.0e-201	1125	
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAV03463	1711.00	3300.56	7.1e-176	1125	
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV03465	1701.00	3281.12	8.7e-175	1125	
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ31056	1037.00	1990.60	6.6e-103	1077	
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ31058	1037.00	1976.54	4.0e-102	3831	
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ31060	1037.00	1975.56	4.5e-102	4185	
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ31065	1037.00	1975.44	4.6e-102	4231	
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ31062	1037.00	1975.37	4.7e-102	4257	
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ31064	1035.50	1973.23	6.1e-102	3969	
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ31063	1033.50	1969.51	9.1e-102	3909	
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ31093	1033.00	1982.79	1.8e-102	1080	
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ19461	1033.00	1982.79	1.8e-102	1080	
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ37399	1013.00	1940.24	4.2e-100	1503	
/SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAZ54416	746.50	1420.70	3.7e-71	1703	
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ49048	746.50	1420.70	3.7e-71	1703	
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ31790	745.00	1421.89	3.1e-71	1402	
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ10854	745.50	1418.13	5.1e-71	1802	
/SIDSL/gcgdata/geneseq/geneseq/NA1993.DAT:AAQ51098	742.50	1412.93	9.9e-71	1703	
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAQ09737	740.00	1408.32	1.8e-70	1664	
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ15502	738.50	1406.24	2.3e-70	1543	
/SIDSL/gcgdata/geneseq/geneseq/NA1994.DAT:AAQ69004	725.50	1376.67	1.0e-68	2274	
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ1294	719.00	1364.22	5.1e-68	2236	
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ57661	707.00	1348.29	3.9e-67	1146	
/SIDSL/gcgdata/geneseq/geneseq/NA1994.DAT:AAQ69002	692.50	1314.44	3.0e-65	1910	
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ53006	668.00	1266.37	1.4e-62	1987	
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ49451	651.00	1236.31	6.8e-61	1516	
/SIDSL/gcgdata/geneseq/geneseq/NA1996.DAT:AAZ14527	643.00	1220.46	5.2e-60	1558	
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ41972	638.00	1210.50	1.9e-59	1591	
/SIDSL/gcgdata/geneseq/geneseq/NA1997.DAT:AAZ62461	601.50	1132.85	4.0e-55	2310	
/SIDSL/gcgdata/geneseq/geneseq/NA1997.DAT:AAZ62460	601.50	1132.63	4.1e-55	2967	
/SIDSL/gcgdata/geneseq/geneseq/NA2001.DAT:AAZ58272	598.00	1128.97	6.5e-55	2235	
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ92615	526.50	994.66	2.0e-47	1461	
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ36330	520.50	981.56	1.1e-46	1662	
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ47419	520.50	981.56	1.1e-46	1662	
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ13598	435.00	825.36	5.3e-38	672	
/SIDSL/gcgdata/geneseq/geneseq/NA2001.DAT:AAZ66090	424.00	809.72	3.9e-37	400	
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ69735	380.00	723.09	2.6e-32	441	
/SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAZ76649	360.00	675.43	1.2e-29	974	
/SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ16974	354.00	665.95	4.0e-29	799	

seq_name: /SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT:AAA48751
seq_documentation_block:
ID AAA48751 standard; cDNA; 2060 BP.
XX
AC AAA48751;
XX
DT 08-SEP-2000 (first entry)
XX
DE Human G-alpha-16 nucleotide sequence.
XX
KW Human; G-alpha-16; G protein; cytostatic; hyperproliferative disorder;
KW cancer; inflammation; infection; antisense inhibition; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 220..1344
FT /*tag= a
FT /product= "G-alpha-16"
XX
PN WO200032817-A1.
XX
PD 08-JUN-2000.
XX
PF 25-AUG-1999; 99WO-US19613.
XX
PR 03-DEC-1998; 98US-0205143.
XX (ISIS-) ISIS PHARM INC.
XX
PI Cowser LM;
XX
DR WPI; 2000-412354/35.
DR P-PSDB; AAY99841.
XX
PT A new antisense compound for inhibiting the expression of human
PT G-alpha-16 and treating, preventing or delaying infections,
PT inflammation or hyperproliferative disorders such as cancer -
XX
PS Example 13; Page 82-84; 100pp; English.
XX
CC The present sequence encodes the human G-protein G-alpha-16.
CC G-alpha-16 interacts differentially with several receptor types
CC including members of the opioid and chemokine receptor families.
CC Antisense oligonucleotides that target the present sequence can be
CC used to modulate the expression of G-alpha-16. They may be used to
CC inhibit the expression of G-alpha-16 in human cells and tissues and thus
CC to treat diseases associated with G-alpha-16, such as hyperproliferative
CC disorders, especially cancer. Infections, inflammation or tumour
CC formation can be prevented or delayed. The compounds can be
CC used in research and diagnostics in sandwich and other assays.
XX
SQ Sequence 2060 BP; 415 A; 631 C; 640 G; 374 T; 0 other;

alignment_scores:
Quality: 1960.00 Length: 374
Ratio: 5.241 Gaps: 0
Percent similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-878-801-2 x AAA48751 ..
Align seg 1/1 to: AAA48751 from: 1 to: 2060
1 MetAlaArgSerLeuThrTrpArgCysCysProTrpCysLeuThrGluAs 17


```
alignment_scores:
  Quality: 1947.00      Length: 374
  Ratio: 5.206          Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 99.198

alignment_block:
  US-08-878-801-2 x AAV03464 ..

  Align seg 1/1 to: AAV03464 from: 1 to: 1125

1 MetAlaArgSerLeuThrPArgCysProTrpCysLeuThrGluAs 17
1 ATGGCCCGCTCGCTGACCTGGCGTGTGCGCCCTGGTGCCTGAGCGAGGA 50
17 pGluLysAlaAlaAlaArgValAspGlnGluLeuAsnArgIleLeuLeu 34
17 TCGAGAGCGCGCGCGCGGTGGACAGAGATCAACAGGATCCTCTTGG 100
34 LuGlnLysGlnAspArgGlyGluLeuLysLeuLeuGlyPro 50
101 AGCAGAAGAAGCAGACCGCGGGAGCTTGAAGCTGCTGCTTTTGGGCCCA 150
51 GlyLysSerGlyLysSerThrPheIleLysGlnMetArgIleIleHisG1 67
151 GCGGAGCGGGAAGAGACCTTTCATCAAGCAGATGCGGATCATCCACGG 200
67 yAlaGlyTyrSerGluGluArgLysGlyPheArgProLeuValTyrG 84
201 CGCGGCTACTCGGAGGAGGAGCGAAGGCTTCGCGCCCTGGTGTACC 250
84 InAsnIlePheValSerMetArgAlaMetIleGluAlaMetGluArgLeu 100
251 AGAACATCTTCGTGTCCATCGCGGGCCATGATCGAGCCCATGGAGCGCTG 300
101 GlnIleProPheSerArgProGluSerLysHisAlaSerLeuValMe 117
301 CAGATTCATTCAGAGCGCGGAGAGCAAGCAGCAGCAGCTAGCTGTGCAT 350
117 tSerGlnAspProTyrLysValThrPheGluLysArgTyrAlaAla 134
351 GAGCCAGGACCCCTATAAAGTGACACCGTTCGAGAAGCGCTACGCTGCGG 400
134 LaMetGlnTrpLeuTyrArgAspAlaGlyIleArgAlaCysTyrGluArg 150
401 CCATGAGTGGCTGTGGAGGATGCGGGCATCGGGCTGCTATGAGCGT 450
151 ArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLe 167
451 CGCGGGAATTCCACCTGCTCGATTAGCCGCTGCTACTACCTGTCCACCT 500
167 uGluArgIleThrGluGluGlyTyrValProThrAlaGlnAspValLeu 184
501 GGAGCGCATCACCGAGGAGGCTACGTGCCACAGCTCAGAGCTGCTCC 550
184 rgSerArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGln 200
551 GCACCGCATGCCACCATGGCATCAACGAGTACTGCTTCTCGTGCAG 600
201 LysThrAsnLeuArgIleValAspValGlyGlnLysSerGluArgLy 217
601 AAACCAACCTGCGGATCGTGGAGTGGGGGCGGAGGTCAGAGCGTAA 650
217 sLysTrpIleHisCysPheGluAsnValIleAlaLeuIleTyrLeuAla 234
651 GAAATGGATCATTTGTTTCGAGAAGCGTATCGCCCTCATCTACTGCGCT 700
234 erLeuSerGluTyrAspGlnCysLeuGluGluAsnAsnGlnGluAsnArg 250
701 CACTGAGTATACGACCATGCTGCGTGGAGGAGAACACAGGAGAACCGC 750
251 MetLysGluSerLeuAlaLeuPheGlyThrIleLeuGluLeuProTrpPh 267
```

```
|||||
751 ATGAGAGAGAGCGCTCGCATTTGTTGGGACTATCTCTGGAACCTACCTGGTT 800
267 eLysSerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuGluG 284
|||||
801 CAAAGACATCGCTCATCTCTTCTCAACAAACCGACATCTCTGGAGG 850
284 LuLysIleProThrSerHisLeuAlaThrTyrPheProSerPheGlnGly 300
|||||
851 AGAAATCCCACTCCACCTCCGCTGCTACCTATTTCCCAAGTTCCAGGCG 900
301 ProLysGlnAspAlaGluAlaLysArgPheIleLeuAspMetTyrTh 317
|||||
901 CCTAAGCAGGATGCTGAGGAGCAAGAGGTTTCATCTGGACATGTACAC 950
317 rArgMetTyrThrGlyCysValAspGlyProGluGlySerLysLysGly 334
|||||
951 GAGGATGATACCGGTGCTGGAGGCGCCCGGAGGCGCAAGAAGGCG 1000
334 laArgSerArgArgLeuPheSerHisTyrThrCysAlaThrAspThrGln 350
|||||
1001 CACGATCCGACGCTTTTCAGCCACTACACATGTGCCACAGACACACAG 1050
351 AsnIleArgLysValPheLysAspValArgAspSerValLeuAlaArgTy 367
|||||
1051 AACATCCGCAAGGCTTCAAGGAGCTGCGGAGCTCGGTGCTCGCCCGCTA 1100
367 rLeuAspGluIleAsnLeuLeu 374
|||||
1101 CCTGGAGCAGATCAACCTGCTG 1122

seq_name: /STDS1/gcgdata/geneseq/geneseq/NA2000.DAT.AAA57483

seq_documentation_block:
ID AAA57483 standard; cDNA; 1125 BP.
XX
AC AAA57483;
XX
DT 03-OCT-2000 (first entry)
XX
DE cDNA encoding a rabbit G-protein alpha 16 polypeptide.
XX
KW Rabbit; G-protein alpha 16; signal transduction; immunogen; ss.
XX
OS Oryctolagus cuniculus.
XX
FH key Location/Qualifiers
FT CDS 1..1125
FT /*tag= a
FT /product= "G-protein alpha 16 polypeptide"
XX
FN WO200036920-A1.
XX
PD 29-JUN-2000.
XX
PE 21-DEC-1999; 99WO-US30726.
XX
PR 22-DEC-1998; 98US-0218489.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
PA (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI Ames RS; Feild JA, Testa T;
XX
WPI; 2000-442490/38.
DR P-PSDB; AAY93970.
XX
PT Novel rabbit G-alpha 16 polypeptides for antibody production and for
PT screening antagonists and agonists of G-coupled protein receptor -
XX
PS Claim 2; Page 25-26; 30pp; English.
XX
CC The present sequence encodes a rabbit G-protein alpha 16 polypeptide.
```



```

213 LysSerGluArgLysTyrTrpIleHisCysPheGluAsnValIleAlaLeu 229
   ::::::::::::::::::::
3733 AGGTCCAGAGAGAAGAAATGGATACACTCGCTTTGAAAATGCACCTCAT 3782
   ::::::::::::::::::::
229 uIeTYrLeuAlaSerLeuSerGluTYrASpGlnCYsLeuGluGluAsnA 246
   ::::::::::::::::::::
3783 CATGTTTCATGTCGCGTTAGTAGGAATATGATCAAGTTCTCGTGGAAGTCAG 3832
   ::::::::::::::::::::
246 snGlnGluAsnArgMetLYsGLuSereuAlaLeuPheGlyThrIleLeu 262
   ::::::::::::::::::::
3833 ACAATGAGACCgaatGGAGGAAGCAAGCGTCTCTTTAGACAATATATC 3882
   ::::::::::::::::::::
263 GluLeuProTrpPheLYsSerThrSerValIleLeuPheLeuAsnLYsth 279
   ::::::::::::::::::::
3883 ACATAccCGCTGGTTCcAGAACTcCTCGGTTATTCTCTCTTAACAAGAA 3932
   ::::::::::::::::::::
279 rAsPIleLeuCluGlUySLyleProThrSerHisLeuAlaThrTYrPheP 296
   ::::::::::::::::::::
3933 AGATCTTCAGAGGAGAAATCATGTATTCCCACATCTAGTCGACTACTCC 3982
   ::::::::::::::::::::
296 toSerPheGlnGlyPrOLysGlnAsPaLaGluAlaLalysArgPheIIle 312
   ::::::::::::::::::::
3983 CAGAATATGATGGACCcCAGAGAGATGCCcAGGcAGCcGAGAAATTCATT 4032
   ::::::::::::::::::::
313 LeuAspMetTYrThrArgMetTYrThrGLYCYsValAspGLYProGluCL 329
   ::::::::::::::::::::
4033 CTGAAGATGTYTCGTGGACCTG.....AACCCAGACAG 4064
   ::::::::::::::::::::
329 ySerLYsLYsGLYAlaArgSerArgLeuPheSerHisTYrThrCYsa 346
   ::::::::::::::::::::
4065 TGACAAA.....ATTATCTACTCCCACTTCACGTGGG 4096
   ::::::::::::::::::::
346 laThrAspThrcLnAsnIleArgLYsValPheLYsAspValArgAspSer 362
   ::::::::::::::::::::
4097 CCACAGACACCgAGAAATATCCGCTTGTCTTGTCGCGCTCAAGGACACC 4146
   ::::::::::::::::::::
363 ValLeuAlaArGYrLeuAspGluIleAsnLeu 373
   :::::
4147 ATCTCTCCAAGTTGAACCTGAAGAGCTGCGGCTG 4179
   :::::
seq name : /SIDSI/gcdadata/geneseq/geneseqn/NA1999.DAT:AAZ31065

```

```
seq_name: /SIDS1/gcgdata/geneseq/WA1999.DAT:AAZ31065

seq_documentation_block:
IID_AAZ31065 standard; cDNA; 4231 BP.
XX
XX_AAZ31065;
XX
XX_07-JAN-2000 (first entry)
XX
XX_pmGluR2/CaR*galphai5+3ala linker fusion construct cDNA sequence.
XX
XX_G-protein fusion receptor; CaR; calcium receptor; GluR; head injury;
XX_metabotropic glutamate receptor; GABABR; chimeric receptor; stroke;
XX_gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
XX_spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
XX_Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
XX_cognitive disorder; ss.
XX
XX_Homo sapiens.
XX_OS
XX_WO9951641-Al.
XX_PN
XX_PD
XX_14-OCT-1999.
XX_PF
XX_02-APR-1999; 99WO-US07333.
XX_PP
XX_03-APR-1998; 98US-0080671.
XX_PR
XX_PA
XX_(NPSP-) NPS PHARM INC.
XX
XX_Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
XX_Simin RT;
XX
```

```

13492  TCATATGATGATGCAATTAAGAGTTATGGAATGATCTCTGAATCCAGG 3541
146   tCysTyrGluArgArgGluPheHisLeuLeuAspSerAlaValTyr 162
3542  ANTGTATGATAGACGACGAATATCAATATCTGACTCTACCAATATAC 3591
163   TyrLeuSerHisLeuGluArgIleThrGluGluGlyTyrValProThrAl 179
3592  TATCTTAATGACTTGGACCGGTAGCTACCCCTGCTACCTGCCTACGCA 3641
179   aGlnAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTyrC 196
3642  ACAAGATGCTTGAAGTTCAGTCCGCCACACAGGGATCATCGAATACC 3691
196   ysPheSerValGlnLysThrAsnLeuArgIleValAspValGlyGln 212
3692  CCTTTGACTTCAAAAGTCTATTTTCAGAAATGTCGATAGGGGCCAA 3741
213   LysSerGluArgLysLysTyrPheHisCysPheGluAsnValIleAlaLe 229
3742  AGGTCAGACAGACAAATGATGATACCTGCTTGAATGTCACCTCTAT 3791
229   uIleTyrLeuAlaSerLeuSerGluTyrAspGlnCysLeuGluGluAsnA 246
3792  CATGTTCTAGTACCGCTAGTGAATGATGATCAAGTTCTCTGCGGAGTCA 3841
246   snGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThrIleLeu 262
3842  ACAATGAGAACCGAATGGAGAAAGCAAGGCTCTTTAGAACCAATATC 3891
263   GluLeuProTrpPheLysSerThrSerValIleLeuPheLeuAsnLysTh 279
3892  ACATACCCTCTGGTTCAGAACTCTCGGTTATCTGTCTTAAACAAGAA 3941
279   rAspIleLeuGluGluLysLysThrSerHisLeuAlaThrTyrPheP 296
3942  AGATCTTCTAGAGGAGAAATCATGTATTCATCTAGTCTGACTACTTCC 3991
296   rSerPheGlnGlyProLysGlnAspAlaGluAlaAlaLysArgPheIle 312
3992  CAGAAATATGATGACCCACAGAGATGCCAGGACGCCGAGATTCATT 4041
313   LeuAspMetTyrThrArgMetTyrThrGlyCysValAspGlyProGluG 329
4042  CTGAAGATGTTCTGGACCTG.....AATCCAGACAG 4073
329   ySerLysLysGlyAlaArgSerArgLeuPheSerHisTyrThrCysA 346
4074  TGACAA.....ATTATCTCTCCACTTCACGTGCG 4105
346   lathrAspThrGlnAsnIleArgLysValPheLysAspValArgAspSer 362
4106  CCACAGACCGGAGATATCCGTTTCTGTCCTGCGCCGCAAGGACAC 4155
363   ValLeuAlaArgTyrLeuAspGluIleAsnLeu 373
4156  ATCTCTCAGTTGAACCTGAAGGACTCGGCTG 4188

```

seq_name: /SIDS1/gcgdata/genesecq/NA1999.DAT:AAZ31062

seq_documentation_block:
ID AAZ31062 standard; cDNA; 4257 BP.

XX AC AAZ31062;

XX DT 07-JAN-2000 (first entry)

DE mGluR8/Car*Galphaq15 fusion construct cDNA sequence.

XX G-protein fusion receptor; Car; calcium receptor; GluR; head injury;
KW metabotropic glutamate receptor; GABABR; chimeric receptor; stroke;
KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;

```

KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;  
KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;  
XX cognitive disorder; ss.  
XX Homo sapiens.  
XX WO9951641-A1.  
XX 14-OCT-1999.  
XX 02-APR-1999; 99WO-US07333.  
XX 03-APR-1998; 98US-0080671.  
XX (NPS-P) NPS PHARM INC.  
XX Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;  
XX Simin RT;  
XX WPI; 1999-610995/52.  
XX P-PSDB; AAY49131.  
XX New G-protein fusion receptors and chimeras containing domains from  
XX different receptors, used to screen for modulators, potentially useful  
XX e.g. for treating or preventing stroke or Alzheimer's disease -  
XX Example 1; Fig 11; 255pp; English.  
XX The invention relates to G-protein fusion receptors (I) comprising:  
XX (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and  
XX intracellular (ICD) domains, each chosen independently from a Car  
XX (calcium receptor), GluR (metabotropic glutamate receptor) and GABABR  
XX (gamma-aminobutyric acid receptor); (2) an optional linker attached to  
XX the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the  
XX linker. (I), and recombinant chimeric receptors (CR) without the GP  
XX component, are used to assess function of the various domains and to  
XX identify compounds (e.g. allosteric modulators or antagonists) that act  
XX on these domains. The modulators are potentially useful for treating or  
XX preventing diseases associated with the receptors, e.g. stroke, head or  
XX spinal cord injury, epilepsy, ischaemia, hypoglycaemia, cognitive  
XX disorders and depression. Nucleic acid (II) that encodes (I) is used:  
XX (1) for recombinant production of corresponding proteins; and (2) to  
XX produce cells used in screening for modulators. Use of Car and mGluR  
XX domains allows presentation of GABABR domains, to a binding agent, in a  
XX form more like the natural domain structure compared with use of  
XX incomplete receptors, lacking one or more domains. By shuffling different  
XX domains, agents can be identified that affect particular domains of a  
XX receptor.  
XX SQ Sequence 4257 BP; 1152 A; 1044 C; 1046 G; 1015 T; 0 other;

```

alignment_scores:

Quality: 1037.00 Length: 361
Ratio: 3.515 Gaps: 2
Percent Similarity: 81.717 Percent Identity: 55.402

alignment_block:

US-08-878-801-2 x AAZ31062

Align seg 1/1 to: AAZ31062 from: 1 to: 4257

```

13 CysLeuThrGluAspGluLysAlaAlaAlaArgValAspGlnGluIleAs 29
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3205 TGCTTGACGCGAGGCGCCAGGAAGCCGCGGNTCAACGACGATCGA 3254
29 nArGlleLeuLeuGluGlnLysLysGlnAspArgGlyGluLeuLeuL 46
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3255 GCGGACGCTCCGCGGACACAGCGGAGCCCGCGCGGAGCTCAAGCTGC 3304
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
46 euteLeuGlyProGlyGluSerGlyLysSerTrpPheIleLysGlnMet 62
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Thu Sep 6 15:16:52 2001

3305 TGCTGCTGGCAGGAGAGAGTGGCAAGAGTACGTTTATCAAGCAGATG 3354
63 ArgilelleHisGlyAlaGlyTyrSerGluGluGluArgLysGlyPheAr 79
3355 AGAATCATCCATGGTCCAGTATCTCTGATGAGATAAAGGGGCTTCAC 3404
79 gProLeuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluA 96
3405 CAAGCTGGTGTATCAGAACATCTTCCAGGCCATGCAGGCCATGATCAG 3454
96 laMetGluArgLeuGlnIleProPheSerArgProGluSerLysHisHis 112
3455 CCATGGACACACTCAAGATCCATACAGATGATGACACATTAAGGCTCAT 3504
113 AlaserLeuValMetSerGlnAspProTyrLysValThrPheGluLy 129
3505 GCACANTTAGTTCGAGAGTGTGATGGAGAAGGTGTCTGTTTGGAGAA 3554
129 sArgTyrAlaAlaLaMetGlnTrpLeuTrpArgAspAlaGlyIleArgA 146
3555 TCCATATGATGATCAATAAAGAGTGTGGAATGATCCTGGAATCCAGG 3604
146 laCysTyrGluArgArgGluPheHisLeuLeuAspSerAlaValTyr 162
3605 AATGCTATGATAGACGACGAGATATCAATATCTGACTCTACCAAAATAC 3654
163 TyrLeuSerHisLeuGluArgIleThrGluGluGlyTyrValProThrAl 179
3655 TATCTTAATGACTTGGACCGCTAGTACCTGCTGCTACCTGCTTACGCA 3704
179 aGlnAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTyrC 196
3705 ACAAGATGTCTAGATGTCGAGTCCACACGAGGATCATCGAATACC 3754
196 yPheSerValGlnLysThrAsnLeuArgIleValAspValGlyGlyGln 212
3755 CCTTGGACTTACAAAGTGTCTTTTTCAGAAATGTCGATGAGGGGCCAA 3804
213 LysSerGluArgLysLysTrpIleHisCysPheGluAsnValIleAlaLe 229
3805 AGGTGACGAGAGAAATGGATGACTGCTTGAANAATGTCACCTCTAT 3854
229 uIleTyrLeuAlaSerLysSerGluTyrAspGlnCysLeuGluGluAsnA 246
3855 CATGTTCTAGTACGCTAGTGAATATGATCAAGTCTCTGTTGAGTACG 3904
246 snGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThrIleLeu 262
3905 ACAATGAGAACCAATGGAGGAAGCAAGGCTCTCTTTAGAACAAATATC 3954
263 GluLeuProThrPheLysSerThrSerValIleLeuPheLeuAsnLysTh 279
3955 ACATACCCCTGGTTCAGAACTCCTCGTTATCTGTTCTTAAACAAGAA 4004
279 rAspIleLeuGluGlyLysIleProThrSerHisLeuAlaThrTyrPheP 296
4005 AGATCTTCTAGAGAGAAATCATGTTATCCCATCTAGTCCGACTACTTCC 4054
296 roSerPheGlnGlyProLysGlnAspAlaGluAlaLysArgPheIle 312
4055 CAGAATATGATGGACCCAGCAGAGATCCCGAGGAGCCGAGGAATTCATT 4104
313 LeuAspMetTyrThrArgMetTyrThrGlyCysValAspGlyProGluGl 329
4105 CTGAAGATGTTCTGGGACCTG.....AATCTACTCCCACTTACGTCG 4136
329 ySerLysGlyAlaArgSerArgArgLeuPheSerHisTyrThrCysA 346
4137 TGACAAA.....ATTATCTACTCCCACTTACGTCG 4168
346 laThrAspThrGlnAsnIleArgLysValPheLysAspValArgAspSer 362
4169 CCACAGACACCGAGAATATCGGCTTTGCTTTTGTGCGCGTCAAGGACAC 4218

363 ValLeuAlaArgTyrLeuAspGluIleAsnLeu 373
4219 ATCTCCCACTTGAACCTGAGACCTGCGGTCTG 4251
seq_name: /SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ31064

seq_documentation_block:
ID AAZ31064 standard; cDNA: 3969 BP.
XX AAZ31064;
XX 07-JAN-2000 (first entry)
XX GABA-BR1a*Gqo5 fusion construct cDNA sequence.

G-protein fusion receptor; CaR: calcium receptor; GluR: head injury;
metabotropic glutamate receptor; GABABR: chimeric receptor; stroke;
gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
KW cognitive disorder; ss.

XX Homo sapiens.
XX PN WO9951641-A1.
XX 14-OCT-1999.

XX 02-APR-1999; 99WO-US07333.

XX 03-APR-1998; 98US-0080671.

XX (NPSP-) NPS PHARM INC.

XX Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
PI Shimin RT;

XX WPI; 1999-610995/52.
XX P-PSDB; AAY49133.

New G-protein fusion receptors and chimeras containing domains from
different receptors, used to screen for modulators, potentially useful
e.g. for treating or preventing stroke or Alzheimer's disease -

XX Disclosure; Fig 13; 255pp; English.

XX The invention relates to G-protein fusion receptors (I) comprising:
CC (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and
CC intracellular (ICD) domains, each chosen independently from a CaR
CC (calcium receptor), GluR (metabotropic glutamate receptor) and GABABR
CC (gamma-aminobutyric acid receptor); (2) an optional linker attached to
CC the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the
CC linker. (I), and recombinant chimeric receptors (CR) without the GP
CC component, are used to assess function of the various domains and to
CC identify compounds (e.g. allosteric modulators or antagonists) that act
CC on these domains. The modulators are potentially useful for treating or
CC preventing diseases associated with the receptors, e.g. stroke, head or
CC spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anoxia,
CC Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive
CC disorders and depression. Nucleic acid (II) that encodes (I) is used:
CC (1) for recombinant production of corresponding proteins; and (2) to
CC produce cells used in screening for modulators. Use of CaR and mGluR
CC domains allows presentation of GABABR domains, to a binding agent, in a
CC form more like the natural domain structure compared with use of
CC incomplete receptors, lacking one or more domains. By shuffling different
CC domains, agents can be identified that affect particular domains of a
CC receptor.

XX Sequence 3969 BP; 945 A; 1079 C; 1065 G; 880 T; 0 other;

alignment_scores:

Quality: 1035.50 Length: 361
 Ratio: 3.510 Gaps: 3
 Percent Similarity: 81.717 Percent Identity: 55.402
 alignment_block:
 US-08-878-801-2 x AA231064 ..
 Align seg 1/1 to: AA231064 from: 1 to: 3969
 13 CysLeuThrGluAspGluLysAlaAlaAlaArgValAspGlnGluLeuAs 29
 2920 TGCCTGAGCGAGGAGCCCAAGGAGGCGCGGCGGATCAACACGAGATCGA 2969
 29 nArgIleLeuLeuGluGlnLysLysGlnAspArgGlyGluLeuLysLeuL 46
 2970 GCGGAGCTCCGCGAGGACAAAGCGGAGCGCGCGCGGAGCTCAAGCTGC 3019
 46 euLeuLeuGlyProGlyGlySerGlyLysSerThrPheIleLysGlnMet 62
 3020 TGCCTGCTGGGACGAGAGAGTGGCAAGAGTACGTTTATCAAGCAGATG 3069
 63 ArgIleHisGlyAlaGlyTyrSerGluGluGluArgLysGlyPheAr 79
 3070 AGAATCATCCATGGTGCAGATACCTGATGAAGATAAAGGGCTTCAC 3119
 79 gProLeuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluA 96
 3120 CAAGCTGGTGTATCAGAACATCTTCAGGCCATGAGCGGATGATCAGAG 3169
 96 laMetGluArgLeuGlnIleProPheSerArgProGluSerLysHisHis 112
 3170 CCATGGACACACTCAAGATCCCTACAGTATGAGACACAAATAGGCTCAT 3219
 113 AlaSerLeuValMetSerGlnAspProTyrLysValThrPheGluLys 129
 3220 CCACAAATAGTTCGAGAGTGTATGAGAGGTTGTCTGCTTTTGAGAA 3269
 129 sArgTyrAlaAlaMetGlnTrpLeuTrpArgAspAlaGlyIleArgA 146
 3270 TCCATGTGTATGATGATAAAGAGTGTATGGAATGATCTCGAATCCAGG 3319
 146 laCysTyrGluArgArgGluPheHisLeuLeuAspSerAlaValTyr 162
 3320 AATGCTATGATGAGCAGGAGAAATATCAATATCTGACTCTACCAATAC 3369
 163 TyrLeuSerHisLeuGluArgIleThrGluGlyTyrValProThrAl 179
 3370 TATCTTATGACTTGGACCGGTAGCTGACCTGCCTACCTGCCTACGCA 3419
 179 aGlnAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTyrC 196
 3420 ACAAGATGTGCTTAGAGTTCAGTCCGCCACACAGGATCATCGAATAC 3469
 196 ysPheSerValGlnLysThrAsnLeuArgIleValAspValGlyGln 212
 3470 CCTTTGACTTACAAAGTGTATTTTCAGATGTCGATGAGGGGCCAA 3519
 213 LysSerGluArgLysLysTyrPheHisCysPheGluAsnValIleAlaLe 229
 3520 AGGTCAGAGAGAGAAATGATGATGCTGCTTCAAAATGTCACCTCTAT 3569
 229 uIleTyrLeuAlaSerLeuSerGluTyrAspGlnCysLeuGluGluAsnA 246
 3570 CATGTTTCTAGTACGCTTAAAGTAAATGATGATGATGCTGCTGAGGTCAG 3619
 246 snGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThrIleLeu 262
 3620 ACAATGAGAACCAATGAGGAAGCAAGGCTCTCTTTAGAACCAATATC 3669
 263 GluLeuProTrpPheLysSerThrSerValIleLeuPheLeuAsnLysTh 279
 3670 ACATACCCCTGGTTCAGAACTCCTCGGTATTCTTCTTAAACAAGAA 3719

279 rAspIleLeuGluGlnLysIleProThrSerHisLeuAlaThrTyrPheP 296
 3720 AGATCTTCTAGAGGAGAAATCATGTATTCCTATCTAGTCGACTACTTCC 3769
 296 roSerPheGlnGlyProLysGlnAspAlaGluAlaLysArgPheIle 312
 3770 CAGATATGATGAGCCAGAGAGATGCCAGCAGCCCGAGAAATTCATT 3819
 313 LeuAspMetTyrThrArgMetTyrThrGlyCysValAspGlyProGluG 329
 3820 CTGAAGATGTCGTGGACCTG.....AACCCAGACAG 3851
 329 ySerLysLysGlyAlaArgSerArgArgLeuPheSerHisTyrThrCysA 346
 3852 TGACAAA.....ATTATCTCTCCACTTCACGTGCG 3883
 346 laThrAspThrGlnAsnIleArgLysValPheLysAspValArgAspSer 362
 3884 CCACAGACACCGAGAAATATCGCTTTGCTTCTGCTCGCTCAAGGACACC 3933
 363 ValLeuAlaArgTyrLeuAspGluIleAsnLeu 373
 3934 ATCCTC.....CAGTTGAACCTG 3951
 seq_name: /SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT:AAZ31063
 seq_documentation_block:
 ID AAZ31063 standard; cDNA: 3909 BP.
 XX AAZ31063;
 XX 07-JAN-2000 (first entry)
 XX GABA-BR2*Gqo5 fusion construct cDNA sequence.
 XX G-protein fusion receptor; CaR: calcium receptor; GluR: head injury;
 KW metabotropic glutamate receptor; GABAAR: chimeric receptor; stroke;
 KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
 KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
 KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
 KW cognitive disorder; ss.
 OS Homo sapiens.
 XX
 XX WO9951641-A1.
 XX 14-OCT-1999.
 XX 02-APR-1999; 99MO-US07333.
 XX 03-APR-1998; 98US-0080671.
 XX (NPSF-) NPS PHARM INC.
 XX Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
 PI Simin RT;
 XX WPI; 1999-610995/52.
 DR P-PSDB; AAY49132.
 XX
 PT New G-protein fusion receptors and chimeras containing domains from
 PT different receptors, used to screen for modulators, potentially useful
 XX e.g. for treating or preventing stroke or Alzheimer's disease
 PS Disclosure; Fig 13; 255pp; English.
 XX
 CC The invention relates to G-protein fusion receptors (1) comprising:
 CC (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and
 CC intracellular (ICD) domains, each chosen independently from a CaR
 CC (calcium receptor), GluR (metabotropic glutamate receptor) and GABAAR
 CC (gamma-aminobutyric acid receptor); (2) an optional linker attached to
 CC the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the
 CC linker. (1), and recombinant chimeric receptors (CR) without the GP

XX 3000 BB. 1010 B. 1097 C. 989 G: 813 T: 0 other;

Percent Identity: 91.717

US-08-878-801-2 x AAZ31063

Итого по 1-й группе 1/1

|||||::|||
TGGGTGCGCCGCCACCGCCGGATCAACGACGAGATCGA 290

29 nArgIleLeuLeuGluGluGlnLysLysGlnAspArgGlyGluLeuLysLeu

2910 GCGGCAGCTCCGCAGGGACAAGCGGGACGCCCGCCGGGAGCTCAATCTGC

63 ArgTlleHisGlvAlaGlyTyrSerGluGluGluArgLysGlyPheAr 79

JULY AUGUST SEPTEMBER OCTOBER NOVEMBER DECEMBER

[illegible]

96 1aMetGluArgLeuGlnIlePropheserArgProGluSerLysnshnts 112

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

3160 GCACAAATTAGTTCGAGAAGTTGATGTGGAGAAGGTGCTGCTTTTGAGAA 3209

[illegible]

...-Glu-Thr-Ileu-Lys-Ser-Ala-Val-Tyr-162

3260 AATGCTATGATAGACGACGAGAAATATCAATTAATCTGACCTACCAATAC

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z .

179 aGlnAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTyrC 196

THE UNIVERSITY OF CHICAGO

..... DAT:AAZ41093

AAZ41093;

Human C-α1pha-11 encoding cDNA.

	D	X
	S	X
	C	
	*	
	S	
	S	
	S	
	I	
	I	
	A	
	T	
	T	

PN WO9953101-A1.

PD 21-OCF-1999.

[illegible]

PR 28-APR-1998;

LL
 YY
 (1040)

Thu Sep 6 15:16:52 2001

us-08-878-801-2.p2n.rng

XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 1..1080
FT /*tag= a
FT /product= "G-alpha-11"
XX
XX US951455-A.
XX PD 14-SEP-1999.
XX PF 04-DEC-1998; 98US-0205922.
XX PR 04-DEC-1998; 98US-0205922.
XX PA (ISIS-) ISIS PHARM INC.
XX PI Cowser LM;
XX DR WPI; 1999-539140/45.
XX DR P-PSDB; AAY29789.
XX PT Inhibitory antisense compounds useful for the treatment of diseases
XX associated with G-alpha-11
XX PS Example 13; Column 45-48; 38pp; English.
XX CC The present invention describes inhibitory antisense compounds of 8-30
XX nucleotides, targeted to a nucleic acid molecule encoding human
XX G-alpha-11. The present sequence encodes human G-alpha-11. AAZ19468 to
XX CC AAZ19547 represent human G-alpha-11 phosphorothioate antisense
XX CC oligonucleotides given in the present invention. The oligonucleotides
XX may be useful for the treatment of diseases associated with G-alpha-11.
XX SQ Sequence 1080 BP; 246 A; 334 C; 314 G; 186 T; 0 other;

alignment_scores:
Quality: 1033.00 Length: 362
Ratio: 3.550 Gaps: 2
Percent Similarity: 80.387 Percent Identity: 55.525
alignment_block:
US-08-878-801-2 x AAZ19461 ..
Align seg 1/1 to: AAZ19461 from: 1 to: 1080
13 CysLeuThrGluAspGluLysAlaAlaArgValaspGlnGluLeu 29
28 TGCCTGAGCAGTGGTGAAGAGTCCAAAGCGGATCAACGCCGAGATCGA 77
29 nArgileLeuGluGlnLysLysGlnAspArgGlyGluLeuLysLeu 46
78 GAACGAGCTGCGGGGACAAAGCGGACGCCCGCGCGAGCTCAAGCTGC 127
46 euLeuLeuGlyProGlyGluSerGlyLysSerThrPheileLysGlnMet 62
128 TGCTGCTGCGCGCGGAGAGCGGAGGAGAGACGCTTCATCAAGCAGATG 177
63 ArgileLeuHisGlyAlaGlyTyrSerGluGluGluArgLysGlyPheAr 79
178 CGCATCATCCACGCCGCGGCTACTCGGAGGAGGAGGAGGAGCGGCTTCAC 227
79 gProLeuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluA 96
228 CAAGCTCTCTACCAAGACATCTTCACCGCCCATCGAGGCCATGATCGGG 277
96 laMetGluArgLeuGlnIleProPheSerArgProGluSerLysHisHis 112
278 CCATGGAGACGCTCAAGATCTCTACAGAGTACGAGACAGACAAAGCCCAAT 327
113 AlaSerLeuValMetSerGlnAspProTyrLysValThrPheGluLys 129

328 GCCTCTGATCGGGAGGTGGAGAGGTGACCACTTTCGAGCA 377
129 sArgTyrAlaAlaAlaMetGlnTrpLeuTrpArgAspAlaGlyIleArgA 146
378 TCAGTACGTACGTGCTCAAGACCTGTGGGAGGACCGGCGATCCAGG 427
146 laCystTyrGluArgArgGluPheHisLeuLeuAspSerAlaValTyr 162
428 AATGCTACGACCGCGGCGGAGTACCACTCTCCGACTCTCCCAAGTAC 477
163 TyrLeuSerHisLeuGluArgIleThrGluGluGlyTyrValProThrAl 179
478 TACCTGACCGACGCTTGACCGCATCGCCTTGGGCTACCTGCCACCCA 527
179 aGlnAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTyrC 196
528 GCAGGAGCTGCTCGGGTCCGCGTCCGACCCACCGCATCATCGAGTACC 577
196 ysPheSerValGlnLysThrAsnLeuArgIleValAspValGlyGln 212
578 CTTTCGACCTGGAGAACATCATCTTCGGATGGTGGATGTGGGGGCCAG 627
213 LysSerGluArgLysLysTrpIleHisCysPheGluAsnValIleAlaLe 229
628 CGGTGGAGCGGAGAGTGGATCCACTGCTTTGAGAACGCTGACATCCAT 677
229 uIleTyrLeuAlaSerLeuSerGluTyrAspGlnCysLeuGluAsnA 246
678 CATGTTCTTCGTCGCCCTCAGCAATACGACCAAGTCTGTGGAGTCGG 727
246 snGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThrIleLeu 262
728 ACAACGAGAACCGGATGGAGGAGAGCAAAAGCCCTGTCCGGAGCATCATC 777
263 GluLeuProTyrPheLysSerThrSerValIleLeuPheLeuAsnLysTh 279
778 ACCTACCCCTCGTCCAGAACCTCCGTCATCTCTCTCTCAACAAGAA 827
279 rAspIleLeuGluGluLysIleProThrSerHisLeuAlaThrTyrPheP 296
828 GGACCTGCTGGAGGACAAAGATCCTGCTACTCGCACCTGTGGACTACTTC 877
296 roSerPheGlnGlyProLysGlnAspAlaGluAlaAlaLysArgPheIle 312
878 CCAGTTCGATGTTCCCGGAGGAGCGCCAGCGCGCGGAGTTCATC 927
313 LeuAspMetTyrThrArgMetTyrThrGlyCysValaspGlyProGluGl 329
928 CCAGAGATGTTCTGGACCTG.....AACCCCGACAG 959
329 ySerLysLysGlyAlaArgSerArgArgLeuPheSerHisTyrThrCysA 346
960 CGACAAG.....ATCATCTACTCACACTTCACGTGTG 991
346 laThrAspThrGlnAsnIleArgLysValPheLysAspValargaspSer 362
992 CCACCGACACGAGAACATCCGCTTCTGTTCCGCGCGGTGAAGGACACC 1041
363 ValLeuAlaArgTyrLeuAspGluIleAsnLeu 374
1042 ATCTGTCAGCTGAACCTGAAGGAGTACAAATCTGGTC 1077
seq_name: /SIDS1/gcgdata/geneseq/geneseqn/NA2000.DAT:AAZ19461

seq_documentation_block:
ID AAA73799 standard; DNA; 1503 BP.
XX
AC AAA73799;
XX
DT 18-DEC-2000 (first entry)
XX
DE Mouse TC-Galpal4 DNA.

XX Mouse; TC-Galpal14; sensory signal; taste; ds.
 XX Mus sp.
 XX Key Location/Qualifiers
 XX CDS 157...1224
 XX /tag- a
 XX /product- TC-Galpal14
 XX WO200044929-A2.
 XX 03-AUG-2000.
 XX 26-JAN-2000; 2000WO-US02217.
 XX 27-JAN-1999; 99US-0117367.
 XX (REGC) UNIV CALIFORNIA.
 XX zucker CS;
 XX WPI; 2000-499336/44.
 XX P-PSDB; AAB15026.
 XX Assaying for compounds that modulate sensory signalling in taste cells,
 XX by determining interactions between the compounds and a sensory cell
 XX specific G-protein alpha subunit polypeptide -
 XX Disclosure; Page 61; 67pp; English.
 XX The present invention relates to a method for identifying a compound
 XX that modulates sensory signalling in taste cells, comprising contacting
 XX the compound with a sensory cell specific G-protein alpha subunit
 XX polypeptide. A suitable peptide is encoded by the present sequence,
 XX mouse TC-Galpal14 DNA. The compounds identified by the present method
 XX may be used by the food and pharmaceutical industries to customize
 XX taste as additives for food or medicines so that they taste different
 XX when eaten.
 XX Sequence 1503 BP; 396 A; 361 C; 374 G; 372 T; 0 other;
 XX

alignment_scores:
 Quality: 1013.00 Length: 366
 Ratio: 3.517 Gaps: 2
 Percent Similarity: 78.689 Percent Identity: 54.372
 alignment_block:
 US-08-878-801-2 x AAA73799 ..
 Align seg 1/1 to: AAA73799 from: 1 to: 1503
 9 CysCysProTrpCysLeuThrGluAspGluLysAlaAlaAlaArgValAs 25
 166 TGTCTG.....TGTGTGTCTGGAGAGAAAGAGTCTCAGCGCATCAG 209
 25 pGlnGluLeuAsnArgLeuLeuGluGlnLysLysGlnAspArgGlyG 42
 210 CGCGAGATCGAGCGACGTTCTGGCCGACAAAGAGAGTGGGACGCGCGG 259
 42 LuLeuLysLeuLeuLeuGlyProGlyGluSerGlyLysSerThrPhe 58
 260 AGCTCAAGCTGCTGTCTGGGAAACCGGTGAGAGTGGGAAAGACCTTT 309
 59 TleLysGlnMetArgTleLeuHisGlyAlaGlyTyrSerGluGluArg 75
 310 ATCAAGCATGAGGATATCATCGGTCTGGCTACAGTGTAGATGAGATAG 359
 75 GlyGlyPheArgProLeuValTyrGlnAsnIlePheValSerMetArg 92
 360 AAAGGGCTTCAGAAAGCTGGTTTACCAAAACATATTCACGGCCATGCAAG 409

92 laMetIleGluAlaMetGluArgLeuGlnIleIleProPheSerArgProGlu 108
 410 CCATGATCAGACCAATGGATACCTCAGGATACAAATACATGTGTGAGCAG 459
 109 SerLysHisAlaSerLeuValMetSerGlnAspProTyrLysValTh 125
 460 AATAAGGAAATGCCCATCATCAGGAAAGTGAAGTAGACAAGGTCTCAC 509
 125 rThrPheGluLysArgTyrAlaAlaMetGlnTrpLeuTrpArgAsp 142
 510 TGCACCTCTAGACACAGGTGGCCAGCATCAAACAGCTGTGGCTGGATC 559
 142 LagIleAlaArgAlaCysTyrGluArgArgGluPheHisLeuLeuAsp 158
 560 CCGAATCCAGAGTGTACACAGGAGGAGGAGTACCAGCTGTCTCAGAC 609
 159 SerAlaValTyrLeuSerHisLeuGluArgIleThrGluGluGly 175
 610 TCTGCCAAATATTACCTGACGAGCATTTGAGCGTATCCCATGCCCTCTT 659
 175 rValProThrAlaGlnAspValLeuArgSerArgMetProThrThrGly 192
 660 CGTGCCAAACACACAGATGCTTCTGTGTAGTGTGCCACCATCTGGCA 709
 192 leAsnGluTyrCysPheSerValGlnLysThrArgLeuArgIleValAsp 208
 710 TCATAGAATATCCATTCGACCTGGAACCATCATCTCCGATGGTGGAT 759
 209 ValGlyGlyGlnLysSerGluArgLysLysTrpIleHisCysPheGluAs 225
 760 GTTGGTGGCCAGCATCTGAACGACGGAATGGTTTCACTGCTTGGAGAG 809
 225 nValIleAlaLeuIleTyrLeuAlaSerLeuSerGluTyrAspGlnCysL 242
 810 TGTACCTCTCATCTTCTGTGTCTCTGAGTGAATATGACCATCTGTC 859
 242 euGluGluAsnAsnGlnGluAsnArgMetLysLysSerLeuAlaLeuPhe 258
 860 TGCTGCTGCTGACATGAGACCCGATGAGAGAGCAAGCAAGCCCTGTT 909
 259 GlyThrIleLeuGluLeuProTrpPheLysSerThrSerValIleLeuPh 275
 910 AGAACCATCATCCTACCTGCTGTTTCTGAACTCTCCGATGATCTGTT 959
 275 eLeuAsnLysThrAspIleLeuGluGluLysIleProThrSerHisLeuA 292
 960 CTAAACCAAGAGGATCTCTAGAGGAGAAATCCTGACTCTCATCTAA 1009
 292 laThrTyrPheProSerPheGlnGlyProLysGluAspAlaGluAlaA 308
 1010 TTAGCTACTTCCAGAGTACACAGACCAAGCAAGATGTCAAAGCGCC 1059
 309 LysArgPheIleLeuAspMetTyrThrArgMetTyrThrGlyCysValAs 325
 1060 AGGACCTTATCTCTGAGCTGTATCAA..... 1086
 325 pGlyProGluGlySerLysLysGlyAlaArgSerIleArgLeuPheSerH 342
 1087GACCAGATCTCTGACAAAGAGTACGTTATCTATCTCTC 1123
 342 isTyrThrCysAlaThrAspThrGlnAsnIleArgLysValPheLysAsp 358
 1124 ACTTCACTGTGTACAGACACCGAGAAATATCCGCTTTGTTGCTGCT 1173
 359 ValArgAspSerValLeuAlaArgTyrLeuAspGluIleAsnLeuLeu 374
 1174 GTCAAGACACAACTCTACAGCTAAACCTACGGGATTCACCTTTGCTG 1221
 seq_name: /SDSL/gcgdata/geneseq/geneseq/NA1998.DAT:AAV54416
 seq_documentation_block:
 D AAV54416 standard; cDNA; 1703 BP.
 XX

Thu Sep 6 15:16:52 2001

us-08-878-801-2.p2n.rng

Page 18

372 nLeu 373
1169 GCTC 1172


```
|||||
351 GAGCCAGACCCCTATAAAGTGACCGCTTTGAGAAGCGGTACCTCGCG 400
134 laMetGlnTrpLeuTrpArgAspAlaGlyIleArgAlaCysTyrGluArg 150
|||||
401 CCATGACAGTGGCTGTGGAGGATGCCGGCATCCGGGCTGCTATGAGCGT 450
151 ArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLe 167
|||||
451 CGCGGGGAATTCCACCTGCTCGATTACCGCTGTACTACCTGTCCACCT 500
167 uGluArgIleThrGluGluGlyTyrValProThrAlaGlnAspValLeuA 184
|||||
501 GGAGCGCATCACCGAGGAGGCTAGCTCCACACAGCTCAGACGCTGCTCC 550
184 rGserArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGln 200
|||||
551 GCAGCGCATGCCACCATCGGCATCAACAGAGTACTGCTTCTCCGTGCAG 600
201 LysThrAsnLeuArgIleValAspValGlyGlyGlnLysSerGluArgLy 217
|||||
601 AAAACCACTGGGATCGTGAGCTCGGGGCCAGAGTCAGAGCGTAA 650
217 sLysTrpIleHisCysPheGluAsnValIleAlaLeuIleTyrLeuAlas 234
|||||
651 GAATGATCCATGTTTCAGAACGTGATCGCCCTCATCTACTCTGGCCT 700
234 eLysSerGluTyrAspGlnCysLeuGluGluAsnGlnGluAsnArg 250
|||||
701 CACTGAGTAATACGACCATGCTGCTGGAGAGAACCAACAGGAGAACCG 750
251 MeTylsGluSerLeuAlaLeuPheGlyThrIleLeuGluLeuProTrpPh 267
|||||
751 ATGAAGAGAGCTCGCATGTTTGGAGCTATCTGGAACCTACCTGGGTT 800
267 eLysSerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuGluG 284
|||||
801 CAAAGACATCCGCTCATCTCTTCTCACAACCGACATCTCGGAGG 850
284 lLysIleProThrSerHisLeuAlaThrTyrPheProSerPheGlnGly 300
|||||
851 AGAAATCCCACTCCACCTGGCTGGCTACCTATTTCCCACTTCCAGGGC 900
301 ProLysGlnAspAlaGluAlaLysArgPheIleLeuAspMetTyrTh 317
|||||
901 CCTAAGCAGGATGCTGAGGAGCAGCAAGAGGTTTCATCTCGACATGTAC 950
317 rArgMetTyrThrGlyCysValAspGlyProGluGlySerLysLysGlyA 334
|||||
951 GAGGATGTACACCGGCTGCTGGACGCGCCCGAGGAGCAGCAAGAGGCG 1000
334 laArgSerArgArgLeuPheSerHisTyrThrCysAlaThrAspThrGln 350
|||||
1001 CAGGATCCGACGCGCTTTCAGCCACTACATGTGCCACAGACACAG 1050
351 AsnIleArgLysValPheLysAspValArgAspSerValLeuAlaArgTy 367
|||||
1051 ACATCCGCAAGGCTCTCAAGGAGCTGCGGGAGCTCGGTGCTCGCCGCTA 1100
367 rLeuAspGluIleAsnLeuLeu 374
|||||
1101 CTTGGACGAGATCAACTGCTG 1122
```

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:us-09-205-143-1

```
; seq_documentation_block:
; Sequence 1, Application US/09205143
; Patent No. 6107091
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowser
; TITLE OF INVENTION: ANTISENSE MODULATION OF G-ALPHA-16 EXPRESSION
; FILE REFERENCE: RTS-0032
; CURRENT APPLICATION NUMBER: US/09/205.143
```

```
; CURRENT FILING DATE: 1998-12-03
; NUMBER OF SEQ ID NOS: 87
; SEQ ID NO 1
; TYPE: DNA
; LENGTH: 2060
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (220)..(1344)
; US-09-205-143-1

alignment_scores:
  Quality: 1960.00      Length: 374
  Ratio: 5.241         Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-08-878-801-2 x US-09-205-143-1 ..
  Align seg 1/1 to: US-09-205-143-1 from: 1 to: 2060

1 MetAlaArgSerLeuThrTrpArgCysCysProTrpCysLeuThrGluAs 17
220 ATGCCCGCTCGCTGACCTGGCGCTGCTGCCCTGGTGCCTGACGGAGGA 269
17 pGluLysAlaAlaAlaArgValAspGlnGluIleAsnArgIleLeuLeu 34
270 TGAGAAGGCCGCCGCCGGTGGACACAGGAGATCAACAGGATCCTCTGG 319
34 luGlnLysLysGlnAspArgGlyGluLeuLysLeuLeuLeuGlyPro 50
320 AGCAGAAGAAGACAGGACCGCGGGAGCTGAAGCTGCTGCTTTGGGCCCA 369
51 GlyGluSerGlyLysSerThrPheIleLysGlnMetArgIleIleHisG 67
370 GCGGAGAGCGGGAAGACACCTTCATCAAGCAGATCGGATCATCCACGG 419
67 YAlaGlyTyrSerGluGluGluArgLysGlyPheArgProLeuValTyrG 84
420 CGCCGCTACTCTCGGAGGAGGAGCAAGGCTTCGGGCCCTGGTCTACC 469
84 InAsnIlePheValSerMetArgAlaMetIleGluAlaMetGluArgLeu 100
470 AGAACATCTTCGTGTCATCGCGGCCATGATCGAGCCATGGAGCGCTG 519
101 GlnIleProPheSerArgProGluSerLysHisHisAlaSerLeuValMe 117
520 CAGATTCCATTACAGAGCCCGAGAGCAAGCACACACCTAGCTAGCTGGTCA 569
117 tSerGlnAspProTyrLysValThrPheGlyLysArgTyrAlaAlaA 134
570 GAGCCAGGACCCCTATAAAGTGACACAGTTTGA3AAGCGCTACGCTGCGG 619
134 laMetGlnTrpLeuTrpArgAspAlaGlyIleArgAlaCysTyrGluArg 150
620 CCATGAGTGGCTGTGGAGGAGTCCGGCATCCGGCCTGCTATGAGCGT 669
151 ArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLe 167
670 CGCGGGGAATTCCACCTGCTCGATTACCGCTGCTACTACCTGTCCACCT 719
167 uGluArgIleThrGluGluGlyTyrValProThrAlaGlnAspValLeuA 184
720 GGAGCGCATCACCGAGGAGGCTACGTCCCCACAGCTCAGGACGCTGCTCC 769
184 rGserArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGln 200
770 GCAGCGCATGCCACCATCGGCATCAACAGAGTACTGCTTCTCCGTGCAG 819
201 LysThrAsnLeuArgIleValAspValGlyGlyGlnLysSerGluArgLy 217
820 AAAACCACTCGGATCGTGAGCTGCGGGAGCTCGGGGCGAGAAAGTCAGAGCGTAA 869
```

```
217 sLysTrpIleHisCysPheGluAsnValIleAlaLeuIleTyrLeuAlaLas 234
|||||
870 GAAATGGATCCATTTGTCGAGAACGTATGCCCTCATCTACCTGGCCT 919
|||||
234 erLeuSerGluTyrAspGlnCysLeuGluGluAsnGlnGluAsnArg 250
|||||
920 CACTGAGTGAATACGACAGTGCCTGGAGGAGAACCAACAGGAGAACGCG 969
|||||
251 MetLysGluSerLeuAlaLeuPheGlyThrIleLeuGluLeuProTrpPh 267
|||||
970 ATCAAGGAGAGAGCTCGCATTTGTTGGGACTATCTGGAACTACCTGGTT 1019
|||||
267 elySerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuGluG 284
|||||
1020 CAAAACCATCGCTCATCTCTTCTCAACAAACCGACATCCTGGAGG 1069
|||||
284 LuLysIleProThrSerHisLeuAlaThrTyrPheProSerPheGlnGly 300
|||||
1070 AGAAAATCCCACTCCACCTCCGCTGGCTACCTATTTCCCAAGTTCCAGGCG 1119
|||||
301 ProLysGlnAspAlaGluAlaLysArgPheIleLeuAspMetTyrTh 317
|||||
1120 CTTAAGCAGATGCTCAGGAGCCAGAGAGGTTCATCTCGACATGTACAC 1169
|||||
317 rArgMetTyrThrGlyCysValAspGlyProGluGlySerLysGlyA 334
|||||
1170 GAGATGTACACCGGCTGGAGCGCCCGGAGGAGGAGGAGGAGGAGG 1219
|||||
334 laArgSerArgArgLeuPheSerHisTyrThrCysAlaThrAspThrGln 350
|||||
1220 CACGATCCGAGCCCTTTTCAGCCACTACACATGTGCCACAGACACACAG 1269
|||||
351 AsnIleArgLysValPheLysAspValArgAspSerValIleAlaArgTy 367
|||||
1270 AACATCCGAAAGGTCTTCAAGGACGTGCGGGACTCGGTGCTGCCCGCTA 1319
|||||
367 rLeuAspGluIleAsnLeuLeu 374
|||||
1320 CTTGGAGGAGATCAACCTGCTG 1341
|||||
```

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-878-801-3

```
seq_documentation_block:
; Sequence 3, Application US/08878801
; Patent No. 6004808
; GENERAL INFORMATION:
; APPLICANT: Negulescu, Paul
; APPLICANT: Offermans, Stefan
; APPLICANT: Simon, Melvin
; APPLICANT: Zuker, Charles
; TITLE OF INVENTION: PROMISCUOUS G-PROTEINS COMPOSITIONS AND THEIR USE
; FILE REFERENCE: 08366/002001
; CURRENT APPLICATION NUMBER: US/08/878,801
; CURRENT FILING DATE: 1997-06-19
; EARLIER APPLICATION NUMBER: US 60/020,234
; EARLIER FILING DATE: 1996-06-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1125
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1122)
US-08-878-801-3
```

alignment_scores:
Quality: 1705.00 Length: 374
Ratio: 4.638 Gaps: 0
Percent Similarity: 97.861 Percent Identity: 84.759

```
alignment_block:
US-08-878-801-2 x US-08-878-801-3
Align seg 1/1 to: US-08-878-801-3 from: 1 to: 1125
1 MetAlaArgSerLeuThrTrpArgCysCysProTrpCysLeuThrGluAs 17
|||||
1 ATGCCCGGCTCCCTGACTTGGGCTGCTGCTCCCTGGTGCCTGACAGAGGA 50
|||||
17 pGluLysAlaAlaAlaArgValAspGlnGluIleAsnArgIleLeuLeuG 34
|||||
51 GGAAAGACTGCCCCCAAGATCGACAGAGAGATCAACAGGATTTGTGG 100
|||||
34 LuGlnLysLysGlnAspArgGlyGluLeuLysLeuLeuLeuGlyPro 50
|||||
101 AACAGAAAAAACAAGAGCGCGAGGAATTGAAACTCCTGCTGTTGGGCGCT 150
|||||
51 GlyGluSerGlyLysSerThrPheIleIysGlnMetArgIleIleHisG1 67
|||||
151 GGTGAGAGCGGGAAGAGTACGTTTCATCAAGCAGATGCGCATTCACGG 200
|||||
67 yAlaGlyTyrSerGluGluArgLysGlyPheArgProLeuValTyrG 84
|||||
201 TGTGGGCTACTCGGAGGAGGACCGCAGAGCCTTCGGGCTGCTCATCTACC 250
|||||
84 InAsnIlePheValSerMetArgAlaMetIleGluAlaMetGluArgLeu 100
|||||
251 AGAATCATCTTCGTCCTCCATGCGAGCCATGATAGATGCGATGGACCGGCTG 300
|||||
101 GlnIleProPheSerArgProGluSerLysHisHisAlaSerLeuValme 117
|||||
301 CAGATCCCTTCAGCAGGCTTCAGCAGGAGGAGGAGGAGGAGGAGGAGGAG 350
|||||
117 tSerGlnAspProTyrLysValThrPheGluLysArgTyrAlaAlaA 134
|||||
351 GACCCAGGAGACCCCTATAAAGTGAGCAGCATTCGAGAGCCATATGACAGTG 400
|||||
134 laMetGlnTrpLeuTrpArgAspAlaGlyIleArgAlaCysTyrGluArg 150
|||||
401 CCATGCGACTACTGTGGGGGAGCGGCGCATCCGTCGATGCTAGCAGCGA 450
|||||
151 ArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLe 167
|||||
451 AGGCGTGAATTCACCTTCCTGGACTCCGCGGTGTATTACCTGTGCACACT 500
|||||
167 uGluArgIleThrGluGluGlyTyrValProThrAlaGlnAspValLeuA 184
|||||
501 GGAGCGCATATCAGAGGACAGCTACATCCCCACTGCGCAAGACGTGTGTC 550
|||||
184 rGserArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGln 200
|||||
551 GCAGTCGCGATGCCACACAGCAGCATCAATGAGTACTGCTTCTCCGTGAAG 600
|||||
201 LysThrAsnLeuArgIleValAspValGlyGlyGlnLysSerGluArgLy 217
|||||
601 AAAACCAAACTCGCATCTGGATGTTGGTGGCCAGAGGTGAGAGCGGTAG 650
|||||
217 sLysTrpIleHisCysPheGluAsnValIleAlaLeuIleTyrLeuAlaLas 234
|||||
651 GAAATGGATTCATCTGTTTCGAGAACGTGATTCCTCATCTACCTGGCCT 700
|||||
234 erLeuSerGluTyrAspGlnCysLeuGluGluAsnGlnGluAsnArg 250
|||||
701 CCCTGAGCGAGTATGACCAAGTGCCTAGAGGAGAGAGATCAGGAGAACGCG 750
|||||
251 MetLysGluSerLeuAlaLeuPheGlyThrIleLeuGluLeuProTrpPh 267
|||||
751 ATGGAGGAGAGTCTCGCTCTGTTTCAGCAGATCCTTAGAGCTGCCCTGGTT 800
|||||
267 elySerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuGluG 284
|||||
801 CAAGAGCACCTCGGTCATCTCTCTCAACAAGAGGAGGACATCCTGGAAG 850
|||||
```

```

284 luLysIleProThrSerHisLeuAlaThrTyPheProSerPheGlnGly 300
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
851 ATAAGATTACACCTCCACCTCCACCTGCCACATCTCCCCAGCTCCAGGGA 900
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
301 ProLysGlnAspAlaGluAlaLysArgPheIleLeuAspMetTyrTh 317
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
901 CCCCGGGGAGACGAGAGCGCGCCCAAGAGCTTCATCTTGGACATGTATGC 950
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
317 rArgMetTyrThrGlyCysValAspGlyProGluGlySerLysLysGlyA 334
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
951 GCGCGTGTACGCGAGCTGCGCAGAGCCCGAGCGGTGGCAGGAAGGCT 1000
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
334 laArgSerArgArgLeuPheSerHisTyrThrCysAlaThrAspThrCln 350
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1001 CCCGCGCGCGCGCTTCTTCCGACACTTCACCTGTGCCACGACACGCAA 1050
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
351 AsnIleArgLysValPheLysAspValArgAspSerValLeuAlaArgTy 367
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1051 AGCGTCCGACGCTGTTCAAGAGCTGCGGAGCTCGGTGCTGGCCCGGTA 1100
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
367 rLeuAspGluIleAsnLeuLeu 374
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
1101 CTTGACGAGATCAACCTGTGT 1122
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::

```

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-09-205-922-1

```

seq_documentation_block:
; Sequence 1, Application US/09205922
; Patent No. 5951455
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowser
; TITLE OF INVENTION: ANTISENSE MODULATION OF G-APLHA-11 EXPRESSION
; FILE REFERENCE: RTS-0030
; CURRENT APPLICATION NUMBER: US/09/205,922
; CURRENT FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 87
; SEQ ID NO 1
; LENGTH: 1080
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1080)
US-09-205-922-1

```

```

alignment_scores:
    Quality: 1033.00      Length: 362
    Ratio: 3.950          Gaps: 2
    Percent Similarity: 80.387      Percent Identity: 55.525

```

```

alignment_block:
US-08-878-801-2 x US-09-205-922-1 ..
Align seg 1/1 to: US-09-205-922-1 from: 1 to: 1080

13 CysLeuThrGluAspGluLysAlaAlaArgValAspGlnGluIleAs 29
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
28 TGCTGAGCGATGAGTGAAGAGGAGTCAAGCGGATCAACGCCGAGATCGA 77
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
29 nArgIleLeuLeuGlnLysLysGlnAspArgGlyGluLeuLeuL 46
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
78 GAAGCAGCTGCGCGCGGACAAAGCGCGACGCCCGCGGAGCTCAAGCTGC 127
    :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
46 euLeuLeuGlyProGlyGluSerGlyLysSerPheIleLysGlnMet 62
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
128 TGCTGCTCGCGCGGCGGAGCGGAGAGGACGACGTTTCATCAAGCAGATG 177
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
63 ArgIleIleHisLysAlaGlyTyrSerGluGluGluArgLysGlyPheAr 79
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
178 CGCATCATCAACGGCGCGGCTACTCTCGGAGGAGGACAAAGCGCGGCTTAC 227
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::

```

```

79 gProLeuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluA 96
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
228 CAAGCTCGTCTACCAAGAACATCTTCAACCGCATGCAGGCCATGATCCGG 277
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
96 laMetGluArgLeuGlnIleProPheSerArgIletoGluSerLysHis 112
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
278 CCATGGAGACGCTCAAGATCCTCTACAAGTACACGACAAACAGCCCAAT 327
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
113 AlaSerLeuValMetSerGlnAspProTyrLysValThrThrPheGlu 129
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
328 GCGTCTCTGATCCGGGAGCTGAGCTGGAGAACGTGACCACCTCTCGACA 377
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
129 sArgTyAlaAlaAlaMetGlnTrpLeuTrpAlaGaspAlaGlyIleArg 146
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
378 TCAGTACGTCAAGTCCATCAAGACCTCTGTGGGTGAGCCCGGGCATCCAG 427
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
146 laCysTyrGluArgArgGluPheHisLeuLeuAspSerAlaValTyr 162
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
428 AATGCTACGACCGCGAGCGCGAGTACCAAGCTCTCCGACTCTGCCAAGTAC 477
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
163 TyrLeuSerHisLeuGluArgIleThrGluGluGlyTyrValProThrAl 179
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
478 TACCTGACCGACGTTGACCGCATCGCCACCTTGGCTACCTGCCACCCA 527
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
179 aglnAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTyrC 196
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
528 GCAGGACGTGCTGCGGGTCCGCGTCCGCCACCCACCGGCATCATCGATACC 577
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
196 ysPheSerValGlnLysThrAsnLeuArgIleValAspValGlyGlyGln 212
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
578 CTTTCGACCTGGAGAACATCATCTTCCGGATGGTGGATGTGGGGGCCAG 627
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
213 LysSerGluArgLysLysTyrIleHisCysPheGluAsnValIleAlaLe 229
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
628 CGGTCGCGCGGAGGAGTGGATCCACTGCTTTGAGAACGTGACATCCAT 677
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
229 uIleTyrLeuAlaSerLeuSerGluTyrAspGlnCysLeuGluLeuAsnA 246
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
678 CATGTTCTCTGTCGCTCAGCAATACGACCAAGTCTGTGTGGAGTCGG 727
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
246 snGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThrIleLeu 262
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
728 ACAACGAGAACCGGATGGAGGAGAGCAAGCCCTGTTCGGGACCATCATC 777
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
263 GluLeuProTrpPheLysSerThrSerValIleLeuPheLeuAsnLysTh 279
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
778 ACCTACCCCTGGTTCAGAACTCCTCGTCATCTCTTCTCAACAAGAA 827
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
279 rAspIleLeuGluGluLysIleProThrSerHisLeuAlaThrTyrPheP 296
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
828 GGACCTGCTGGAGGACAAAGTCTCTACTCGCACTGTGGTGGACTACTTCC 877
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
296 roSerPheGlnGlyProLysGlnAspAlaGluAlaLysArgPheIle 312
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
878 CCGAGTTTCGATGTTCCCGACGCGGACGCCAGGCGCGCGGAGTTCATC 927
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
313 LeuAspMetTyrThrArgMetTyrThrGlyCysValAspGlyProGluG 329
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
928 CCGAAGATGTTCTGGACCTG.....AACCCCGACAG 959
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
329 ysSerLysLysGlyAlaArgSerArgLeuPheSerHisTyrThrCysA 346
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
960 CGAAG.....ATCATCTCACTACATTCACGTGTG 991
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
346 laThrAspThrGlnAsnIleArgLysValPheLysAspValArgAspSer 362
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
992 CCACCGACGCGGAGAACATCCGTTCTGTGTTCGCGCGCGGAGTTCATC 1041
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
363 ValLeuAlaArgTyrLeuAspGluIleAsnLeuLeu 374
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::
1042 ATCCTGACGCTGAACCTGAAGGAGTACAACTGTGTC 1077
    :::::::::::::: :: :: :: :: :: :: :: :: :: :: :: ::

```

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-07-868-353A-11

[illegible]

```

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-407-804-20
seq_documentation_block:
; Sequence 20, Application US/08407804
; Patent No. 5817759
; GENERAL INFORMATION:
; APPLICANT: Margolskee, Robert F.
; TITLE OF INVENTION: Gustducin Materials and Methods
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/407,804
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/045,801
; FILING DATE:
; APPLICATION NUMBER: US 07/868/353
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5817759and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1703 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 114..1175
; US-08-407-804-20

alignment_scores:
  Quality: 746.50      Length: 368
  Ratio: 2.893        Gaps: 4
  Percent Similarity: 70.109   Percent Identity: 42.391

alignment_block:
US-08-878-801-2 x US-08-407-804-20 ..

Align seg 1/1 to: US-08-407-804-20 from: 1 to: 1703

14  LeuThrGluAspGluLysAlaAlaAalaArgValAspGlnGluLeuAsnAr 30
126 ATTAGTTTCAGAGACCAAGGAGTCCAGCAAAAGGTCCTCAAGAACTGGAGAA 175
30  gileLeuGluGlnLysLysGlnAspArgGlyLeuLeuLeuLeuL 47
176 GAAGCTTCAGAGATGCTGAACAGAGATGCAAGAACTGTGAAGTTGCTGC 225
47  euLeuGlyProGlyGluSerGlyLysSerThrPheLeuGlnMetArg 63
226 TATTAGGAGCAGGGAATCAGAAAGAGTACTATTGTATTAACAATGAG 275

```

```

64  IleIleHisGlyAlaGlyTyrSerGluGluGluArgLysGlyPheArgPr 80
276 ATCATCCCAAGAAGTGGTTACAGTAACAAGATTCGCATGGAGTTTAAAGC 325
80  oLeuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluAlam 97
326 AGTGGTTTACAGTAACAAGTGGTTCAGTCCATCCCGGCCATTGTGAAGCCA 375
97  eGluArgLeuGlnIleProPheSerArgProIleSerLysHisHisAla 113
376 TGACTACACTAGGATGATTGATTCATCCGGAAGTAGAGAGACCAA 425
114 SerLeuValMetSerGlnAspProTyrLysValThrThrPheGlu.... 128
426 CAACTGCTCTCTCCATG.....GCAAACACACTAGGAAGATGG 463
129 .....LysArgTyrAlaAlaAlaMetG..nTrrPleTrrArgAspA 142
464 TGACATGACGCCCTCAGTTGGCTGAAATAATTAACGTCCTGTGGCGCATC 513
142 laGlyIleArgAlaCysTyrGluArgArgIlePheHisLeuLeuAsp 158
514 CAGGAATTCAGCTGCTTCGAAAGGCATCTGAATACAGCTCAATGAC 563
159 SerAlaValTyrTyrLeuSerHisLeuGluArgIleThrGluGluGly 175
564 TCTGCAGCTTACTACCTTAATGACTTAGATAGACTCACAGCCCTGGGTA 613
175 rValProThrAlaGlnAspValLeuArgSerArgMetProThrThrGlyI 192
614 TGTGCCAAATGAACAGAGCTTCTACATTCCTCCGGTGAACACCTGGTA 663
192 leAsnGluTyrCysPheSerValGlnLysThrAsnLeuArgIleValAsp 208
664 TCATTGAACTCAATTCCTCTTAAAGACTGTAACCTTCAGAAATTTTGT 713
209 ValGlyGlnLysSerGluArgLysLysTrpIleHisCysPheGluAs 225
714 GTAGTGGCCAGAGATCAGAAAGAAAGAAATGATCCACTGCTTTTGAAG 763
225 nValIleAlaLeuIleTyrLeuAlaSerLeuSerGluTyrAspGlnCysL 242
764 AGTGACCTGCATTTATATTTTGTGCGAGCCCTAAATGCTACGACATGGTAC 813
242 euGluGluAsnAsnGlnGluAsnArgMetLysIleSerLeuAlaLeuPhe 258
814 TTGTAGAAGATGAAGAGGTGAACAGAAATGCATGAAGTCTTCACTCTTC 863
259 GlyThrIleLeuGluLeuProTrrPheLysSerThrSerValIleLeuPh 275
864 AACAGCATCTGTAATACAGATATTTTGCACACACCTCCATTGTTCTGTT 913
275 eLeuAsnLysThrAspIleLeuGluLysIleProThrSerHisLeuA 292
914 TCTTAACAAGAAAGATCTCTTCCAGGAGAAAGTGCACCAAGGTGCACCTCA 963
292 laThrTyrPheProSerPheGlnGlyProLysIleAspAlaGluAlaAla 308
964 GCATCTGTTTCCCAAGATACACTGGACCAAAATACATTCGAAGATGCAGGG 1013
309 .....LysArgPheIleLeuAspMetTyrThrArgMetTyrThrGl 322
1014 AACTACATCAAGAACCAGTCTCTAGACCTGAATCTTAAATA..... 1052
322 yCysValAspGlyProGluGlySerLysLysGlyAlaArgSerArgArgL 339
1053 .....AAAGAAGATAAGGAAA 1068
339 euPheSerHisTyrThrCysAlaThrAspThrIleAsnIleArgLysVal 355
1069 TCTATTCTCAGATCACCTGCCTACTGCACACA:AAAACGTCAAATTCGTG 1118

```



```

211 GlyGlnLysSerGluArgLysLysTrpIleHisCysPheGluAsnValI 227
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
619 GGTCAAGATCTGACGGGAAGAAGTGGATTCACTTCTTCCGAAGGAGTCA 668
227 eAlaLeuIleValLeuAlaSerLeuSerGluTrpAspGlnCysLeuGlu 244
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
669 GCGCATCATCTTCTGTGTAGCAGTCAAGTCACTACCACTGGTTAGCTG 718
244 LuAsnAsnGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThr 260
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
719 AGATGAAGAATAATGAACCGATGCATGAAGCATGAAATTTGTTGACAGC 768
261 IleLeuGluLeuProTrpPheLysSerThrSerValIleLeuPheLeuAs 277
|||:|:|:|||||:|||||:|||||:|||||:|||||:|||||:
769 ATATGTGAACAACAAGTGGTTTACAGATACATCCATTACTATTTTCTAAA 818
277 nlysthrAspIleLeuGluGluLysIleProThrSerHisLeuAlaThr 294
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
819 CAAGAAGGATCTCTTTGAAGAAAAAATCAAAAGAGCCCTCTCACTATAT 868
294 yzPheProSerPheGlnGlyProLysGlnAspAlaGluAlaAlaLysArg 310
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
869 GCTATCCAGATATCCAGATCAACATATGAAGAGGCAGCT..... 912
311 PheIleLeuAspMetTyrThrArgMetTyrThrGlyCysValAspGlyPr 327
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
913 .....GCATATATTCATGTCAGTTT..... 933
327 oGluGlySerLysLysGlyAlaArgSerArgArgLeuPheSerHisTyrT 344
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
934 :GAAGACCTCAATAAAAGAAAGGACACAAAGGAAATATACACCACTTCA 982
344 hrCysAlaThrAspThrGlnAsnIleArgLysValPheLysAspValArg 360
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
983 CATGTGCCACAGATCTACTAGAAATGTGCAGTTGTTTTTGATGCTGTAA 1032
361 AspSerValLeuAlaArgTyrIleuAspGluIleAsnLeu 373
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1033 GATGTCTCATATAAAAAATAATCTAAAGAGATTGTGCTCTC 1071

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-09-339-993-1

seq_documentation_block:
; Sequence 1, Application US/09339993A
; Patent No. 6040179
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF G-ALPHA-T2 EXPRESSION
; FILE REFERENCE: RTS-0064
; CURRENT APPLICATION NUMBER: US/09/339,993A
; CURRENT FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 1664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (78)..(1145)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 1105
; OTHER INFORMATION: unknown
; US-09-339-993-1

alignment_scores:
Quality: 740.00 Length: 366
Ratio: 2.879 Gaps: 4
Percent Similarity: 70.219 Percent Identity: 42.896

alignment_block:

```

us-08-878-801-2.p2n.rni

[illegible]

```

000 .....5 .....
978 GCCAGC...TATCCAGAGTAAGTTT..... 1000
324 lAspGlyProGluGlySerLysLysGlyAlaAlaI.gSerArqLeuPheS 341
002 .....GAGGACCTGAATTAAGCGCAAGG/CACCAAGAGAGATCTACA 1041
341 exHisTyrThrCysAlaThrAspThrGlnAsnIleAargLysValPheLys 357
042 CGCACTTCACGTGCGCCACCGACACCAAGAAGACTGCGAGTTGCGTTTGAC 1091
358 AspValArgAspSerValLeuAlaIArgTyrLeuAspGluIleAsnLeu 373
092 GCGGTCCACGATGNCATCATCAAGAACAAGCTCAAGGACATGCGGCGCTC 1139
name: /cgn2_6/ptodata/2/1na/6A_COMB.s;q:US-09-339-775-1

documentation_block:
sequence 1, Application US/09339775
ent No. 6063626
ERAL INFORMATION:
PLICANT: Lex M. CowserT
TITLE OF INVENTION: ANTISENSE MODULATION OF G-ALPHA-I3 EXPRES-
LE REFERENCE: RUS-0069
CURRENT APPLICATION NUMBER: US/09/339,775
CURRENT FILING DATE: 1999-06-24
NUMBER OF SEQ ID NOS: 47
) ID NO 1
LENGTH: 1543
YPE: DNA
ORGANISM: Homo sapiens
EATURE:
NAME/KEY: CDS
LOCATION: (9)..(1073)
-339-775-1

ment_scores:
Quality: 738.50 length: 363
Ratio: 2.885 Gaps: 4
ent Similarity: 70.523 Percent Identity: 43.802

ment_block:
08-878-801-2 x US-09-339-775-1 ..
on seg 1/1 to: US-09-339-775-1 from 1 to: 1543

14 LeuThrGluAspGluLysAlaAlaIArgVa.AspGlnGluIleAsnAr 30
||||: .....||||:||||:||||:||||:||||:||||:||||:||||:
21 TTGAGCGCCGACACAGACGCGCGCAGTGCGAGGAGGAGCAAGATGATCGACCG 70
30 gIleLeuGluGlnLysLysGlnAspArGg.YgluLeuLysLeuLeuL 47
71 CAACCTTACGGAGAGCGGGGAAAGGCGGCCCAAGAAGTGAAGCTGCTGTC 120
47 eutLeuGlyProGlygluSerGlyLysSerThrPheIleLysGluMetArg 63
121 TACTCGGTGCTGGAGAATCTGGTAAAGACCAATGTGAACACAGATGANA 170
64 IleIleHisGlyAlaGlyTyrSerGluGluGlnArgLysGlyPheArGPr 80
171 ATCATTTATCAGGATGCTATTTCAGAGGATGAATGTAAACAATATAAAGT 220
80 oLeuValTyrGlnAsnIlePheValSerMeta:GAlaMetIleGluAlaM 97
221 AGTTGTCTACAGCAATACTACTACATCCATCATTTGCAATCAATAAGCCA 270
97 etGluArgLeuGlnIleProPhe.....SerArqProGluSerLys 110
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

comment_scores:
  Quality: 738.50
  Ratio: 2.885
  Percent Similarity: 70.523
  Percent Identity: 43.802
  Gaps: 4
  Length: 363

```

ment block:

UN sec 1/1 to: US-09-339-775-1 from 1 to: 1543

[illegible]

```
271 TGGACGGCTAAAGATTGACTTTGGGAAGCTCCAGGCGAGATGATGCC 320
111 HisHisAlaSerLeuValMetSerGlnAspProTyrLysValThrPh 127
      ::::: ::::: ::::: ::::: ::::: :::::
321 CGCAATTATTGTTTGTAGCTGGCAGTGCTGAAGAAGAGTCATGATCC 370
127 eGluLysArgTyrAlaAlaAlaMetGlnTrpLeuTrpArgAspAlaGlyI 144
      ||| ||| ::::: ::::: ::::: ::::: :::::
371 AGAA.....CTAGCAGGAGTGATTAACGGTTATGGCGAGATGGTGGG 414
144 leargAlaCysTyrGluArgArgGluPheHisLeuLeuLeuAspSerAla 160
      ::::: ||| ::::: ||| ::::: ||| ::::: |||
415 TACAAGTTCTCTCAGCAGATCCAGGAATATCAGCTCAATGATTCGTGCT 464
161 ValTyrTyrLeuSerHisLeuGluArgIleThrGluGluGlyTyrValPr 177
      ||| ||| ||| ::::: ||| ::::: ||| ::::: |||
465 TCATATTATCTAAATGATCTGGATAGATATCCAGTCTAACTACATCC 514
177 oThrAlaGlnAspValLeuArgSerArgMetProThrThrGlyIleAsnG 194
      ||| ||| ::::: ::::: ::::: ::::: :::::
515 AACTCAGCAAGATGTTCTCGGACGAGAGTGAAGACCACAGGCATTGTAG 564
194 lutyrcysPheSerValGlnLysThrAsnLeuArgIleValAspValGly 210
      ||| ||| ::::: ::::: ::::: ::::: :::::
565 AAACATATTTCACCTTCACAGACTATATCTCAAGATGTTGTATGATAGT 614
211 GlyGlnLysSerGluArgLysLysTyrIleHisCysPheGluAsnValI 227
      ||| ||| ::::: ||| ||| ||| ||| ||| ||| ||| |||
615 GGCCAAAGATCAGAACCAAAAAGTGGATTCACGCTTTTGGAGGAGTGAC 664
227 eAlaLeuIleTyrLeuAlaSerLeuSerGluTyrAspGlnCysLeuGlu 244
      ||| ||| ::::: ::::: ||| ||| ||| ||| ||| |||
665 AGCAATTATCTCTGTGTGCGCCCTCAGTGATATGACCTTGTCTGTGCTG 714
244 luanAsnGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThr 260
      ||| ||| ::::: ||| ||| ||| ||| ||| ||| |||
715 AGGACGAGGAGATGAACGATGATGAAAGCATGAAACTGTTTGCAGC 764
261 IleLeuGluLeuProTyrPheLysSerThrSerValIleLeuPheLeuAs 277
      ||| ||| ::::: ||| ||| ||| ||| ||| ||| ||| |||
765 ATTTGTAATTAACAATGGTTTACAGAACTTCAATCATCTCTCTCCCTAA 814
277 nLysThrAspIleLeuGluGluTyrIleProThrSerHisLeuAlaThr 294
      ||| ||| ::::: ||| ||| ||| ||| ||| ||| ||| |||
815 CAAGAAAGACCTTTTTCAGGAAAAATAAAGAGGAGTCCGTTAACTATCT 864
294 yrPheProSerPheGlnGlyProLysGlnAspAlaGluAlaLysArg 310
      ::::: ||| ||| ::::: ||| ||| ||| ||| ||| |||
865 GTTATCCAGATACACAGTTCCTCAATACATATGAAGAGGAGCT..... 908
311 PheIleLeuAspMetTyrThrArgMetTyrThrGlyCysValAspGlyPr 327
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
909 .....GCCTATATTCAATGCCAGTTT..... 929
327 oGluGlySerLysLysGlyAlaArgSerArgArgLeuPheSerHisTyr 344
      ||| ||| ::::: ::::: ::::: ::::: :::::
930 .GAAGATCTGAACAGAGAAAGATACCAAGAGATCTATACTCACITCA 978
344 hrCysAlaThrAspThrGlnAsnIleArgLysValPheLysAspValArg 360
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
979 CCTGTGCACAGACACGCAAGAAATGTGAGTTTGTGATGCTGTTACA 1028
361 AspSerValLeuAlaArgTyrLeuAspGluIleAsnLeu 373
      ||| ||| ::::: ||| ||| ||| ||| ||| ||| |||
1029 GATGTCATCATTAACACACTTAAGGAATGTGGACTT 1067
```

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:us-08-371-930-28

seq documentation_block:
; Sequence 28, Application US/08371930
; Patent No. 5578451
; GENERAL INFORMATION:
; APPLICANT: Nishimoto, Ikuo

```
? TITLE OF INVENTION: ALZHEIMER'S DISEASE THERAPEUTICS
? NUMBER OF SEQUENCES: 30
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Fish & Richardson
? STREET: 225 Franklin Street
? CITY: Boston
? STATE: Massachusetts
? COUNTRY: U.S.A.
? ZIP: 02110-2804
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3 5" Diskette, 1.44 Mb
? COMPUTER: IBM PS/2 Model 50Z or 55SX
? OPERATING SYSTEM: MS-DOS (Version 5.0)
? SOFTWARE: WordPerfect (Version 5.1)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/371,930
? FILING DATE:
? CLASSIFICATION: 436
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/019,208
? FILING DATE: February 18, 1993
? ATTORNEY/AGENT INFORMATION:
? NAME: Clark, Paul T.
? REGISTRATION NUMBER: 30,162
? REFERENCE/DOCKET NUMBER: 00786/154001
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617) 542-5070
? TELEFAX: (617) 542-8906
? TELEX: 200154
? INFORMATION FOR SEQ ID NO: 28:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2274
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? US-08-371-930-28
```

alignment_scores:
Quality: 731.50 Length: 360
Ratio: 2.950 Gaps: 3
Percent Similarity: 68.889 Percent Identity: 43.611

alignment_block:

US-08-878-801-2 x US-08-371-930-28 ..

Align seg 1/1 to: US-08-371-930-28 from: 1 to: 2274

```
14 LeuThrGluAspGluLysAlaAlaArgValAspGlnGluIleAsnAr 30
      |||::: ::::: |||::: ::::: |||:::
36 CTGAGCGCAGAGAGAGAGCGCCCTCGAGCGGAGCAAGCGGATTGAGAA 85
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
30 gIleLeuLeuGluGlnLysLysGlnAspArgGlyGluLeuLysLeuL 47
      : ||| |||::: ::::: |||::: |||::: |||::: |||
86 AAACCTCAAGAGAGATGGCATCGCGCCCAAGACGCTGAATTTACTCC 135
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
47 euLeuGlyProGlyGluSerGlyLysSerThrPheIleLysGlnMetArg 63
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
136 TGTCTGGGGCTGGAGATCAGAAAAAGACCATTGTGAAGCAGATGAAG 185
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
64 IleIleHisGlyAlaIaGlyTyrSerGluGluArgLysGlyPheArgPr 80
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
186 ATCATCCATGAAGATGGCTTCTCTGGGGAAGACGCTGAAGCAGTACA 235
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
80 oLeuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluAla 97
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
236 TGTGTCTACAGCAACACCATCCAGTCTCTGTGGCGCCATTGTCCGGCCA 285
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
97 etGluArgLeuGlnIleProPheSerArgProGluSerLysHisAla 113
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
286 TGGACACTTTGGCGGTGGAGTGTGTGACAAGGAGGAGGACGAGCTCC 335
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
114 SerLeuValMetSer.....GlnAspProTyrLysValTh 125
```

```

336 AAGATGGTGTGTGACGTGGTGTGATGGAAGAC.....ACTGA 376
125 rThrPheGluLysArgTyrAlaAlaAlaMetGlnTrpLeuTrpArgAspA 142
137 ACCTGTTCTGTGACAACTCTTCTGCTGATGCGACCTGTGGGGCAGCT 426
142 laGlyIleArgAlaCysTyrGluArgArgGluPheHisLeuLeuAsp 158
147 CGGGATCCAGGAGTCTTCAACCGATCTCGGGAGTATCAGCTCAATGAC 476
159 SerAlaValTyrTyrLeuSerHisLeuGluArgIleThrCluGluGly 175
477 TCTGCCAATACTACTCTGACACCGCTGGATCGGATGGACCGGTGACTA 526
175 rValProThrAlaGlnAspValLeuArgSerArgMetProThrThrGly 192
527 CAGCCCACTGACGAGCATCTCCGACCGAGTCAAAACCAACAGTGGCA 576
192 leAsnGluTyrCysPheSerValGlnLysThrAsnLeuArgIleValAsp 208
577 TCTAGAAACCCACTTCACTTCAAGAACCTCCACTTCAGGCTGTTGAC 626
209 valGlyGlnLysSerGluArgLysLysTyrIleHisCysPheGluAs 225
627 GTGGGGGCGAGGATCTGAACCAAGAGTGGATCCACTGCTTTGAGGA 676
225 valIleAlaLeuIleTyrLeuAlaSerLeuSerGluTyrAspGlnCysL 242
677 TGTACGGGCGCATCTCTGTGTCTGCACTCAGCGGCTATGACAGGTGC 726
242 euGluGlnAsnAsnGlnGluAsnArgMetLysGluSerLeuAlaLeuPhe 258
727 TCCACGAGGACGAAACCAACGACCGCATGCACGAATCCCTGAAGCTTTC 776
259 GlyThrIleLeuGluLeuProTrpPheLysSerThrSerValIleLeuPh 275
777 GACAGCATCTGCACAAACAAAGTGTTCACAGACACATCTATATCTGTT 826
275 eteAsnLysThrAspIleGluGluLysIleProThrSerHisLeuA 292
827 TCTCAACAGAGGACATATTGGAGGAGAGATCAAGAAGTCCCACTCA 876
292 laThrTyrPheProSerPheGlnGlyProLysGlnAspAlaGluAlaAla 308
877 CCATCTGCTTCTCTGAATACACACGCCCGGCTTTCACAGAGTGTG 926
309 LysArgPheIleLeuAspMetTyrThrArgMetTyrThrGlyCysValas 325
927 GCTCAC.....ATCCA 937
325 pGlyProGluGlySerLysLysGlyAlaArgSerArgLeuPheSerH 342
938 AGGCGAGTATGAGAGTAAAGTAAGTCAAGTCAACAAGGAGTCTACAGCC 987
342 isTyrThrCysAlaThrAspThrGlnAsnIleArgLysValPheLysAsp 358
988 ATGTCACTGTGCGACGACACCAACATCAATCTCTTTTGTATGCC 1037
359 ValArgAspSerValLeuAlaArgTyrLeu 368
1038 GTGACAGATGTCATCATCGCCAAACCTA 1067
seq_name: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US94-01712-28
seq_documentation_block:
; Sequence 28, Application PC/TUS9401712
; GENERAL INFORMATION:
; APPLICANT: Nishimoto, Ikuo
; TITLE OF INVENTION: ALZHEIMER'S DISEASE THERAPEUTICS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson

```

```

; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01712
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/019,208
; FILING DATE: February 18, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00786/154001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2274
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; PCT-US94-01712-28

alignment_scores:
Quality: 731.50 Length: 360
Ratio: 2.950 Gaps: 3
Percent Similarity: 68.889 Percent Identity: 43.611

alignment_block:
US-08-878-801-2 x PCT-US94-01712-28 ..
Align seg 1/1 to: PCT-US94-01712-28 from: 1 to: 2274
14 LeuThrGluAspGluLysAlaAlaAlaArgValAspGlnGluIleAsnAr 30
136 CTGAGCGCAGAGAGAGAGAGCGCCCTCGAGCGGAGCAAGCGGATGAGAA 85
30 gileLeuLeuGluGlnLysLysGlnAspArgGlyGluLeuLeuLeuL 47
86 AAACCTCAAAGAAGATGGCATCAGCGCGCCCAAGACGTAATTAATCTCC 135
47 euLeuGlyProGlyGluSerGlyLysSerThrPheIleLysGlnMetArg 63
136 TGTCTGGGGCTGGAGAATCAGGAAAGAACACCATTTGTGAAGCAGATCAAG 185
64 IleIleHisGlyAlaGlyTyrSerGluGluArgLysGlyPheArgPr 80
186 ATCATCATCAAGATGCTTCTCTGGGAAGACGTGAAGCAGTACAAGCC 235
80 oleuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluAla 97
236 TGTGGTCTATACCAACACCATCAGTCTCTGGCGCCATGTCCGGGCCA 285
97 etCluArgLeuGlnIleProPheSerArgProGluSerLysHisAla 113
286 TGGACACTTTGGGCGTGGAGTATGTTGACAAAGGAGAGGAGCGGACTCC 335
114 SerLeuValMetSer.....GlnAspProTyrLysValTh 125
336 AAGATGGTGTGTGACGTGTGAGTCTGTATGGAACACTGA 376
125 rThrPheGluLysArgTyrAlaAlaAlaMetGlnTrpLeuTrpArgAspA 142

```



```

725  TCCACGAGGAGCAAAACACGACCGCATGACGAGTCTCTCATGCTCTTC 774
259  GlyThrIleLeuGluLeuProTirPheLysSerThrSerValIleLeuPh 275
      ::::: ::::: ::::: ::::: ::::: ::::: :::::
775  GACTCCATCTGTAACAACAAGTTTTTTCATTCATACCTCCATCCTCTT 824
275  eLeuAsnIlyThrAspIleLeuGluGluLysIleProThrSerHisLeuA 292
      ||||| ::::: ::::: ::::: ::::: |||||
825  CCTCAACAAGAAAGACCTCTTTGGCGAGAAGATTAAAGATGCACCCCTGA 874
292  IaThrTyrPheProSerPheGlnGlyProLysGlnAspAlaGluAlaAa 308
      :: ||||| ::::: ||| ::::: :::::
875  CCATCTGCTTTCCGATATCCAGGCTCCACCACTATGAAGATGAGCT 924
309  LysArgPheIleLeuAspMetTyrThrArgMetTyrThrGlyCysValAs 325
      ::::: :::::
925  ..GCCTACATCCAAACACAGTTTGAA..... 948
325  pGlyProGluGlySerLysGlyAlaArgSerArgArgLeuPhSerH 342
      ||||| ::::: ::::: ::::: ::::: |
949  .....AGCAAAACCCGCTCACCACCAAGAAATTTACTGTGC 985
342  IsTyrThrCysAlaThrAspThrGlnAsnIleArgLysValPheLysAsp 358
      || ||||| ::::: ::::: ::::: |||||
986  ACATGACTTGTGCCACAGACACAGTAATATCCAGGTGGTATTTCGCGCC 1035
359  ValArgAspSerValLeuAlaArgTyrLeu 368
1036  GTCCACCGCATCATCTTTGGCAATCTC 1065

```


us-08-878-801-2.p2n.rst

Thu Sep 6 15:16:55 2001

```

|||||
401 GCGAGAGCGGGAAGACACCTTCATCAAGCAGATGCGGATCATCCACGG 450
67 yAlaGlyTyrSerGluGluGluArgLysGlyPheArgProLeuValTyrG 84
|||||
451 GCGCGCTACTCGGAGGAGGAGCGAAGGCTTCGGGCCCTGGTCTACCG 500
84 InAsnIlePheValSerMetArgAlaMetIleGluAlaMetGluArgLeu 100
|||||
501 AGAACATCTTCGTGCTCATCGGGCCATGATCGAGGCCATGAGCGGCTG 550
101 GlnIleProPheSerArgProGluSerLysHisHisAlaSerLeuValMet 117
551 CAGATTCCATTTCAGAGCCCGGAGCAAGCACCGCTAGCTGGTGCAT 600
117 tSerGlnAspProTyrLysValThrPheGluLysArgTyrAlaAlaA 134
|||||
601 GAGCCAGGAGCCCTATTAAGTGACCAAGTTTGAGAAAGCGCTACGCTG 650
134 laMetGlnTrpLeuTrpArgAlaGlyIleArgAlaCysTyrGluArg 150
651 CCATGCACTGCTGCTGAGGAGTCCGCGCATCCGGGCTGCTATGACGCT 700
151 ArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLe 167
701 CGCGGGGAATTCACCTGCTGCTGATTCAGCCGCTGCTACCTGCTCCAC 750
167 uGluArgIleThrGluGluGlyTyrValProThrAlaGlnAspValLeu 184
751 GAGCGGCATACCGAGGAGGCTAGTCCACACAGCTCAGGACGCTGCC 800
184 tSerArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGln 200
801 GAGCGGCATACCGAGGAGGCTAGTCCACACAGCTCAGGACGCTGCC 850
201 LysThrAsnLeuArgIleValAspValGlyGlnLysSerGluArgLys 217
851 AAAC..AACCTGGGATCTGGACGCTCGGGGC...AGATCAGACGCTA 896
217 sLysTrpIle 220
897 AAAATGGATC 906

seq_name: gb_est80:BE915352
seq_documentation_block: 939 bp mRNA EST 29-SEP-2000
LOCUS BE915352 601667635F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3967497 5',
DEFINITION mRNA sequence.
ACCESSION BE915352
VERSION BE915352.1 GI:10414894
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 939)
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM9142 row: f column: 10
High quality sequence stop: 689.
Location/Qualifiers
1..939
FEATURES
Source
|||||
401 GCGAGAGCGGGAAGACACCTTCATCAAGCAGATGCGGATCATCCACGG 450
67 yAlaGlyTyrSerGluGluGluArgLysGlyPheArgProLeuValTyrG 84
|||||
451 GCGCGCTACTCGGAGGAGGAGCGAAGGCTTCGGGCCCTGGTCTACCG 500
84 InAsnIlePheValSerMetArgAlaMetIleGluAlaMetGluArgLeu 100
|||||
501 AGAACATCTTCGTGCTCATCGGGCCATGATCGAGGCCATGAGCGGCTG 550
101 GlnIleProPheSerArgProGluSerLysHisHisAlaSerLeuValMet 117
551 CAGATTCCATTTCAGAGCCCGGAGCAAGCACCGCTAGCTGGTGCAT 600
117 tSerGlnAspProTyrLysValThrPheGluLysArgTyrAlaAlaA 134
|||||
601 GAGCCAGGAGCCCTATTAAGTGACCAAGTTTGAGAAAGCGCTACGCTG 650
134 laMetGlnTrpLeuTrpArgAlaGlyIleArgAlaCysTyrGluArg 150
651 CCATGCACTGCTGCTGAGGAGTCCGCGCATCCGGGCTGCTATGACGCT 700
151 ArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLe 167
701 CGCGGGGAATTCACCTGCTGCTGATTCAGCCGCTGCTACCTGCTCCAC 750
167 uGluArgIleThrGluGluGlyTyrValProThrAlaGlnAspValLeu 184
751 GAGCGGCATACCGAGGAGGCTAGTCCACACAGCTCAGGACGCTGCC 800
184 tSerArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGln 200
801 GAGCGGCATACCGAGGAGGCTAGTCCACACAGCTCAGGACGCTGCC 850
201 LysThrAsnLeuArgIleValAspValGlyGlnLysSerGluArgLys 217
851 AAAC..AACCTGGGATCTGGACGCTCGGGGC...AGATCAGACGCTA 896
217 sLysTrpIle 220
897 AAAATGGATC 906

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_lib="NCI_CGAP_Mam1"
/dev_stage="3 months, biopsy sample"
/lab_host="DH10B"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site: 1; Sali:
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: G.Libert Smith, NIH"
BASE COUNT 237 a 243 c 105 t 1 others
ORIGIN
alignment_scores: 991.50 length: 251
Quality: 4.219 Gaps: 5
Ratio: 93.625 Percent Identity: 80.876
Percent Similarity:
alignment_block:
US-08-878-801-2 x BE915352 from: 1 to: 939
Align seg 1/1 to: BE915352
1 MetAlaArgSerLeuThrTrpArgCysCysP:ofrCysLeuThrGluAs 17
2 ATGGCCCGGCTCCCTGACTTGGGGCTGCTGTC:CTGGTGGCTGACACAGGA 51
17 pGluLysAlaAlaAlaArgValAspGlnGluIleAsnArgIleLeuLeu 33
52 GAGAGAGACTGCGCCAGAAATCGACGAGGATCAACAGGATTTGTG... 98
34 GluGlnLysGlnAspArgGlyGluLeuLysLeuLeuLeuGlyPr 50
99 GAACAGAAAACACAGAGCGGAGGAATTTGAAACTCTCTGCTGTG.GGGCC 147
50 oGlyGluSerGlyLysSerThrPheIleLysGlnMetArgIleIleHisG 67
148 TGGTGAGAGCGGGAAGAGTACGTTTCATCAACAGATGCGCATCATTCACG 197
67 lyAlaGlyTyrSerGluGluGluArgLysGlyPheArgProLeuValTyr 83
198 GTGTGGGCTACTCGAGGAGGAGCGGAGAG:CTTCCGGGCTGCTCATCTAC 247
84 GlnAsnIlePheValSerMetArgAlaMetIleGluAlaMetGluArgLe 100
248 CAGAACATCTTCGTCTCCATCGAGGCCCATGATAGATGCGATGGACCGCT 297
100 uGlnIleProPheSerArgProGluSerLysHisHisAlaSerLeuValM 117
298 GCAGATCCCTTCAGCAGGCTGACAGCAAGCGAGCGGCGG:ATCCGTGCATCGTACGAGCG 347
117 etSerGlnAspProTyrLysValThrThrGluLysArgTyrAlaAla 133
348 TGACCCAGAGCCCTATAAGTGCAGCACA:TCGAGAAGCCATATGCGAGT 397
134 AlaMetGlnTrpLeuTrpArgAlaGlyIleArgAlaCysTyrGluAr 150
398 GCCATGCACTGCTGCTGCGGAGCGGCGG:ATCCGTGCATCGTACGAGCG 447
150 gArgArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisL 167
448 AAGCGGTGAATTCACCTTCTGGACTCGGCGGTGCTATTACCTGCACACC 497
167 euGluArgIleThrGluGluGlyTyrValProThrAlaGlnAspValLeu 183
498 TGGAGCGCATATCAGAGGACAGCTACATC:CCCAGTCCGCAAGAGCGTGTG 547
184 ArgSerArgMetProThrThrGlyIleAsnGluTyrCysPheSerValG 200
548 CGCAGTCGATGCGCCACACAGGCGATCA:ATGAGTACTGCTTCTCGGTGAA 597

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1071YC21"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/notes="vector: pCMVSPORT 6; site_1: NotI; 1st strand cDNA
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by

```



```

239 TGCAGTGGCTGTGGAGGAGTCCCGCATCCGGCTCGTCTATGAGCGTCGG 190
152 ArgGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLeuG1 168
|||||
189 CGGAATTCACCTGCTGCTGATTCAGCGCTGACTACCTGCTCCACCTGGA 140
|||||
168 uArgIleThrGluGluGlyTyrValProThrAlaGlnAspValLeuArgS 185
|||||
139 GCGCATCACCAGGAGGCGTACGTCCTCCACAGCTCAGGACGTGCTCCGCA 90
|||||
185 erArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGlnLys 201
|||||
89 GCGCATGCCCCACCTACGTCATCAACGAGTACTGCTTCTCCGTCGCCAAA 40
|||||
202 ThrAsnLeuArg 205
|||||
39 ACCAACCTGGCA 28

seq_name: gb_est46:AW368893

seq_documentation_block:
LOCUS AW368893 470 bp mRNA EST 04-FEB-2000
DEFINITION IL2-HT0199-241099-013-F04 HT0199 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW368893
VERSION AW368893.1 GI:6873543
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 470)
HCGP http://www.ludwig.org.br/ORESTES.
TITLES The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=IL2&t2=IL2-HT0199-241099-013-F04&t3=1999-10-24&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 7
High quality sequence stop: 436.
Location/Qualifiers
1..470
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0199"
/dev_stage="Adult"
/notice="Organ: head_neck; Vector: puc18; Site_1: Smal;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 121 a 138 c 121 g 90 t
ORIGIN

alignment_scores:
Quality: 782.00 Length: 151
Ratio: 5.213 Gaps: 0
Percent Similarity: 99.338 Percent Identity: 98.675

alignment_block:
US-08-878-801-2 x AW368893

```

Align seg 1/1 to: AW368893 from: 1 to: 470

```

185 SerArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGlnLy 201
:::|||||
18 GCTCGAATGCCACCACTGGCATCAACGAGTACTGCTTCTCCGTGCAGAA 67
|||||
201 sThrAsnLeuArgIleValAspValGlyGlnLysSerGluArgLysL 218
|||||
68 AACCAACCTTCGGATCGTGGACGTGGGGGCCAGAGTCAGAGCGTAAGA 117
|||||
218 ysTrpLleHisCysPheGluAsnValIleAlaLeuIleTyrLeuAlaSer 234
|||||
118 AATGATCCATTTGTCGAGAACGTGATCGCCCTCATCTACCTAGCCTCA 167
|||||
235 LeuSerGluTyrAspGlnCysLeuGluGluAsnGlnGluAsnArgMe 251
|||||
168 CTGAGTGAATACGACCACTGGCTGGAGGAGAACCAACGAGGAGACCGCAT 217
|||||
251 tLysGluSerLeuAlaLeuPheGlyThrIleLeuGluLeuProTrpPheL 268
|||||
218 GAAGGAGAGCGCTCGCATTTGTTGGACTATCTCTGGAACCTACCTGGTCA 267
|||||
268 ysSerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuGluGlu 284
|||||
268 AAGCACATCGTCTATCTCTTCTCAACAAACCGACATCTCTGGAGGAG 317
|||||
285 LysIleProThrSerHisLeuAlaThrTyrPheProSerPheGlnGlyPr 301
|||||
318 AAAATCCCACTCCCACTCCCACTGCTGCTATTTTCCCACTTTCCAGGGCCC 367
|||||
301 oLysGlnAspAlaGluAlaAlaLysAqPheIleLeuAspMetTyrThrA 318
|||||
368 TAGCAGGATGCTGAGGAGCAGAGAGTTTCATCTGGACATGTACAGCA 417
|||||
318 rgMetTyrThrGlyCysValAspGlyProGluGlySerLysLysGlyAla 334
|||||
418 GGATGTACACCGTGTGCGTGGCGTCCCGAGGCGCAGCAAGAGGGCGCA 467
|||||
335 Arg 335
|||
468 CGA 470

```

seq_name: gb_est87:BF443684

seq_documentation_block: 489 bp mRNA EST 01-DEC-2000
LOCUS BF443684
DEFINITION 261269 MARC 2F1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BF443684
VERSION BF443684.1 GI:11503776
KEYWORDS EST.
SOURCE pig.

ORGANISM

Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

REFERENCE

1 (bases 1 to 489)
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keeler,J.W.

TITLE

Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine

JOURNAL

Unpublished (2000)
Contact: Smith TPL

COMMENT

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt.trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers

FORWARD: AGGAAACACGCTATGACCAT

```

BACKWARD: GTTTTCCAGTCACGACG
Plate: 94 Row: 1 Column: 13
Seq primer: ATTTAGTGACACTATAG.
FEATURES
    Source
        location/Qualifiers
            1..489
            /organism="Sus scrofa"
            /db_xref="taxon:9823"
            /clone_lib="MARC 2Flg"
            /lab_type="pooled"
            /lab_host="Bhi08"
            /note="Vector: pcMV SPORT6; Site_1: XbaI; Site_2: XhoI;
            Library made from pooled tissue from testis, ovary,
            endometrium, hypothalamus, pituitary, and placenta."
            117 a 145 c 93 t 2 others
BASE COUNT
ORIGIN
alignment_scores:
    Quality: 778.00 Length: 162
    Ratio: 4.955 Gaps: 1
    Percent Similarity: 96.914 Percent Identity: 93.210
alignment_block:
    US-08-878-801-2 x BF443684 ..
Align seg 1/1 to: BF443684 from: 1 to: 489
205 ArgIleValAspValGlyGlnLysSerGluArgLysLysTrpIleHi 221
5 AGGATCGTGGAGTTGGGGGCGCATGAGTCAGAACCAAGAAGTGGATCCA 54
221 sCysPheGluAsnValIleAlaLeuIleTyrLeuAlaSerLeuSerGluT 238
55 CTGCTTCTACAAATGTGATCGGCCCTCATCTAGCCTGCGTGAAGTGAAT 104
238 YrAspGlnCysLeuGluGluAsnAsnGlnGluAsnArgMethylGluSer 254
105 ATGACCAAGTGTCTGGAGGAAACCAACCAAGAGACCGAATGAAGGAAAGC 154
255 LeuAlaLeuPheGlyThrIleLeuGluLeuLeuProTrrPheLysSerThrse 271
155 CTGGCCCTGTGGCGCACCATCTTGGAACTGGCCCTGGTTCAAAGAGCACTTC 204
271 rValIleLeuPheLeuAsnLysThrAspIleLeuGluGluLysIleProT 288
205 TGTCTATCTCTCTCTCAACAAACCGACATCTGGAGGAGAGATCCGCCA 254
288 hrSerHisLeuAlaThrTyrPheProSerPheGlnGlyProLysGlnAsp 304
255 CTCTCCACCTGGCTACCTACCTTCCCACTTCCAGGGGCCCGAAGCAGGAC 304
305 AlaGluAlaAlaLysArgPheIleLeuAspMetTyrThrArgMetTyrTh 321
305 CGCGAGGCGAGCAAGAGAGTTTCATCTTGACATGTACACCAAGATGTACGC 354
321 rGlyCysValAspGlyProGluGlySerLysLysGlyAlaArgSerArgA 338
355 CGGCTCGCTGTGATGCGGCGGATGAGGCGAGGAAAGGCCCGCGCTCCCGCC 404
338 rg.LeuPheSerHisTyrThrCysAlaThrAspThrGlnAsnIleArgly 354
405 GNCCTTTTACGCCACTACACGTGTGCCACGACACACAGAATATCCGCAA 454
354 sValPheLysAspValArgAspSerValLeuAla 365
455 GGTCTTCAAGGACGTCGCGGACTCGGTCTCGGCC 488
seq_name: gb_est98:BG258764
seq_documentation_block:
LOCUS BG258764 1153 bp mRNA EST
DEFINITION 602378022F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4509052 5',
mRNA sequence.
13-FEB-2001

```



```

1162 GATGTCATCATTAACAACTTAAGGAATGTGGCTT 1200
seq_name: gb_est78:BE740136
seq_documentation_block:
LOCUS BE740136 768 bp mRNA EST 15-SEP-2000
DEFINITION 601595077F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3949146 5',
mRNA sequence.
ACCESSION BE740136
VERSION BE740136.1 GI:10154128
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 768)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLC813 row: 1 column: 19
High quality sequence stop: 735.
FEATURES
Location/Qualifiers
1..768
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3949146"
/clone_lib="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="PH10B (phage-resistant)"
/note="Organ: ovary; Vector: pOTB7; Site:1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5',
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 171 a 241 c 236 g 130 t
ORIGIN

```

```

alignment_scores:
Quality: 730.50 Length: 258
Ratio: 3.462 Gaps: 5
Percent Similarity: 81.783 Percent Identity: 58.527
alignment_block:
US-08-878-801-2 x BE740136 ..
Align seg 1/1 to: BE740136 from: 1 to: 768
36 LysLysGlnAspArgGlyGluLeuLysLeuLeuLeuGlyProGlyG1 52
|||||:||||| 11 AAGCGGACCGCGCGCGAGCTCAAGCTGCTGCTCGCGCGCGA 60
52 uSerGlyLysSerThrPheIleLysGlnMetArgIleIleHisGlyAlaG 69
61 GAGCGGGAAGAGCAGCTTATCAGCAGATGCGCATCATCCACGCGCGG 110
69 lYrSerGluGluGluArgLysGlyPheArgProLeuValTyrGlnAsn 85
|||||:||||| 111 GCTACTCGGAGGAGGAGCGCGGCTTACCAAGCTCGTCTACCAAGAAC 160
86 IlePheValSerMetArgAlaMetIleGluAlaMetGluArgLeuGlnI1 102

```

```

|||||:||||| 161 ATCTTCACCGCCATGAGCCCATGATCCGGGCCATGGACAGCTCAAGAT 210
102 eProPheSerArgProGluSerLysHisAlaSerLeuValMetSerG 119
|||||:||||| 211 CCTCTACAGTACGAGCAGACAAAGCCCAATGC3CTCCTGATCCGGGAGG 260
119 lnaAsProTyrLysValThrPheGluLysArgTyrAlaAlaAlaMet 135
|||||:||||| 261 TGGACGTGGAGAGGTGACCACTTCGAGCATCAGTACGTGAGTCCATC 310
136 GlnTyrLeuTyrArgAspAlaGlyIleArgAlaLysTyrGluArgArg 152
|||||:||||| 311 AGACCCGTGGAGAGCCCGGCATCCAGGAATGCTACGACCGCAGGCG 360
152 gGluPheHisLeuLeuAspSerAlaValTyrTyrLeuSerHisLeuGlu 169
|||||:||||| 361 CAGTATACCAAGCTCTCCGACTCTGCAAGTACTA:CTGACCGCAGTTGACC 410
169 rGileThrGluGluGlyTyrValProThrAlaGlnAspValLeuArgSer 185
|||||:||||| 411 GCATCGCCACCTTGGGCTACCTGCCACCCAGCAGGAGTGTCTGCGGTC 460
186 ArgMetProThrThrGlyIleAsnGluTyrCysPheSerValGlnLysTh 202
|||||:||||| 461 CGCGTCCCGACCGCGCATCATCGAGTACCTT:TCGACCTGGAGAACAT 510
202 rAsnLeuArgIleValAspValGlyGlnLysSerGluArgLysLysT 219
|||||:||||| 511 CATCTTCGGGATGTTGGATGTGGGGGCCAGCG:TCGAGCGGAGGAAGT 560
219 rPileHisCysPheGluAsnValIleAlaLeuIeTyrLeuAlaSerLeu 235
|||||:||||| 561 GGATCCACTGCTTTCAGAACGTGACATCCATCA:GTTTCTGTCGCCCTC 610
236 SerGluTyrAspGlnCysLeuGluGluAsnAsnGln:GluAsnArgMetL 252
|||||:||||| 611 AGCAATACGACCAAGCTCTGCGAGTCCCGACACGAGAACCGGATGG 660
252 ysGluSerLeuAlaLeuPheGlyThrIleLeuGlu.....LeuProTrp 266
|||||:||||| 661 AGGAGAGCAAGCCCTGTTCGGACCATCATC:TACCCTGGTTCCG... 707
267 PheLysSerThrSerValIleLeuPheLeuAsnLysThrAspIleLeuG1 283
|||||:||||| 708 .....AATCTCTCCGTCATCTCTTC...TCACAAGAAGGACTGCTGGA 748
283 uGluLysIleProThrSerHis 290
|||||:||||| 749 GGACAGA...TCTGTTTCGCAC 767
seq_name: gb_est15:AI040021
seq_documentation_block:
LOCUS AI040021 890 bp mRNA EST 24-SEP-1998
DEFINITION ox97g03.x1 Soares_senescent.fibroblasts.NbHSF Homo sapiens cDNA
clone IMAGE:1664308 3', similar to gb:M69013 GUANINE
NUCLEOTIDE-BINDING PROTEIN G(Y), ALPHA SUBUNIT (HUMAN);, mRNA
sequence.
ACCESSION AI040021
VERSION AI040021.1 GI:3279215
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 890)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov

```

Thu Sep 6 15:16:55 2001

us-08-878-801-2.p2n.rst

This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 1033 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 497.

FEATURES

source

```

1. .890
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1664308"
/clone_lib="Soares_senescent_fibroblasts_NbHSP"
/tissue_type="senescent_fibroblast"
/lab_host="Dh10B (ampicillin resistant)"
/notes="Vector: p773D (Pharmacia) with a modified
polylinker V.TYPE: phagemid; Site_1: Not I; Site_2: Eco RI
; 1st strand cDNA was primed with a Not I - oligo(df)
primer [5',
TGTACCACTCTGAAGTGGGAGCGCGCATTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified p773 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT      199 a      272 c      263 g      153 t      3 others
ORIGIN

```

```

alignment_scores:
  Quality: 730.00      Length: 284
  Ratio: 3.333        Gaps: 2
  Percent similarity: 77.113 Percent Identity: 53.873

```

alignment_block:

US-08-878-801-2 x AI040021

Align seg 1/1 to: AI040021 from: 1 to: 890

```

13 CysLeuThrGluAspGluLysAlaAlaAArgValAspGlnLileAs 29
32 TGCCTGAGCATGAGTGAAGGAGTCCAAAGCGGATCAACGCCGAGATCGA 81
29 nArgIleLeuGluGlnLysLysGlnAspArgGlyGluLeuLysLeuL 46
82 GAAGCAGCTCGCGCGGCGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGG 131
46 euLeuLeuGlyProGlySerGlyLysSerThrPheLysGlnMet 62
132 TGCTGCTCGGCGGCGGCGGAGAGCGGGAAGACGACGTTTCATCAAGCAGATG 181
63 ArgIleLeuGlyAlaGlyTyrSerGluGluGluArgLysGlyPheAr 79
182 CGCATCATCCAGCGCGGCGGCTACTCGGAGGAGGAGCAAGCGCGCTTCAC 231
79 gProLeuValTyrGlnAsnIlePheValSerMetArgAlaMetIleGluA 96
232 CAAGCTCGCTACCAAGACATCTTCAACCGCATCGAGGCCATGATCGGG 281
96 laMetGluArgLeuGlnIleProPheSerArgProGluSerLysHis 112
282 CCATGAGACGCTCAAGATCTCTCAACAGTACGAGCAAGAGGTGACCACTCGAGCA 381
113 AlaSerLeuValMetSerGluAspProTyrLysValThrPheGluLy 129
332 GCGCTCTGATCCGGAGGTGACCTGGAGAGGTGACCACTCGAGCA 381
129 sArgTyrAlaAlaMetGlnTrpLeuTrpArgAspAlaGlyIleArgA 146
382 TCAGTACGTACGTGCGCATCAAGACCTCTGTGGAGGAGCGCGCATCCAGG 431
146 laCysTyrGluArgArgGluPheHisLeuLeuAspSerAlaValTyr 162
432 AATGCTACGACCGCGGCGGAGTACCACTCTCGGACTCTGCCAAGTAC 481

```

```

163 TyrLeuSerHisLeuGluArgIleThrGluGluGlyTyrValProThrAl 179
|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:
482 TACTGACGACGAGTTCACCGCATCGNCACCTTGGGCTACCTGCCACCCA 531
179 aGlnAspValLeuArgSerArgMetProThrThrGlyIleAsnGluTyrC 196
|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:
532 GCAGGACGCTGCTCGGGTCCGCGTCCGCCACCGCATCATCGAGTACC 581
196 ysPheSerValGlnLysThrAsnLeuArgIleValAspValGlyGlyGln 212
|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:
582 CTTTCGACCTGGAGACATCATCTTNCGGATGTTGGATGTGGGGGCCCA 631
213 LysSerGluArgLysTyrPheHisCysPheGluAsnValIleAlaLe 229
|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:
632 CGGTCGCGGAGGAGGAGTGCCTGCTNTGAGAACGTGACATCCAT 681
229 uIleTyrLeuAlaSerLeuSerGluTyrAspGlnCysLeuGluGluAsnA 246
|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:
682 CATGTTTCTCGTCCGCTCAGCGAATACGACCAAGTCTCTGGTGGAGTCG 731
246 snGlnGluAsnArgMetLysGluSerLeuAlaLeuPheGlyThrIleLeu 262
|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:
732 ACCAGGAGAACCGGATGGAGGAGAGCAAGCCCTGTC.CGGGACATCATC 780
263 GluLeuProTyrPheLysSerThrSerValIleLeuPheLeuAsnLysTh 279
|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:
781 ACCTACCCCTGGTCCAGACTCTTCGTCATCTCTTACCAAGGAAG 830
279 rAspIleLeuGluGlyLysIleProThrSerHisLeuAlaThrTyrPheP 296
|||||...:|||||...:|||||...:|||||...:|||||...:|||||...:
831 ACCT...GCTGGAGACAAGATACGACTTGC...ACTGGTGGCTACTTCC 874
296 ro 296
875 CC 876
seq_name: gb_est3:AA175435
seq_documentation_block: 505 bp mRNA EST 16-FEB-1997
LOCUS AA175435 Mus musculus cDNA clone IMAGE:618528
DEFINITION ms87c01.r1 Soares mouse 3NbMS Mouse G protein alpha subunit (MOUSE);,
5', similar to gb:M80632 Mouse G protein alpha subunit (MOUSE);,
mRNA sequence.
ACCESSION AA175435 GI:1756593
VERSION AA175435.1
KEYWORDS house mouse.
SOURCE EST.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 505)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilsson,R. and
Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:379352
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 451.
FEATURES
source
1. .505

```

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_image="618528"
/clone_lib="Scars mouse 3NbMS"
/sex="male"
/tissue_type="Spleen"
/dev_stage="4 weeks"
/lab_host="DH10B"
/notes="vector: p7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTACCAATCTGAAGTGGAGCGCGCTGTCTTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p7T3 vector. RNA
provided by Dr. Bertrand Jordan. Library went through
three rounds of normalization, and was constructed by
Bento Soares and M.Fatima Bonaldo."
BASE COUNT      123 a 143 c 145 g 94 t
ORIGIN

```

```

alignment_scores:
  Quality: 726.50      Length: 160
  Ratio: 4.657        Gaps: 2
  Percent Similarity: 97.500      Percent Identity: 86.875

```

```
alignment_block:
US-08-878-801-2 x AA175435

```

```
Align seg 1/1 to: AA175435 from: 1 to: 505
```

```

96 AlaMetGluArgLeuGlnIleProPheSerArgProGluSerHisH12
15 GCGATGACCGGCTGCAGATCCCTTCAGCGCCGCTGACGAG...CA 61
112 sAlaSerLeu.ValMetSerGlnAspProTyrLysValThrPheGlu 128
62 CGCCAGCCTACGTGATGACCCAGGACCCCTATTAAGTGACCATTCGAG 111
129 LysArgTyrAlaAlaMetGlnTTPLeuTTPArgAspAlaGlyIleAr 145
112 AAGCCATATGCATGGCGCATGCAGTACCTGTGCGGACCGCGCATCCG 161
145 gAlaCysTyrGluArgArgGluPheHisLeuLeuAspSerAlaValT 162
162 TGCATGCTACGAGCGAGGCGTGAATTCACCTTCCTGGACTCCGCGGT 211
162 YrTyrLeuSerHisLeuGluArgIleThrGluGluGlyTyrValProth 178
212 ATTACCTGTACACCTTGGAGCGCATATCAGAGGACAGCTACATCCCA 261
179 AlaGlnAspValLeuArgSerArgMetProThrThrGlyLeuAsnGluT 195
262 GCGCAAGAGGTGCTGCGCATGCGCATGCCACACAGGATCAATGATGA 311
195 rCysPheSerValGlnLysThrAsnLeuArgIleValAspValGlyGly 212
312 CTGCTCTCCGCTGAAGAAACCAAACTCGCATCTGCTGATGTTGTCGC 361
212 LnLysSerGluArgLysTyrIleHisCysPheGluAsnValIleAla 228
362 AGAGGTCAGAGCGTAGGAATGATGATTCACCTGCTCGAGAACGTCAT 411
229 LeuIleTyrLeuAlaSerLeuSerGluTyrAspGlnCysLeuGluGlu 245
412 CTATCTACCTGCGCTCCCTGAGGAGGATGATGACCATGCTGCTGAG 461
245 nAsnGlnGluAsnArgMetLysGluSer 254
462 CATCAGGAGAACCGCATGGAGGAGAGT 489

```

```
seq_name: gb_est14:AA951538
```

```

seq_documentation_block: 717 bp mRNA EST 24-NOV-1998
LOCUS AA951538
DEFINITION LD32085.Sprime LD Drosophila melanogaster embryo pOT2 Drosophila
melanogaster cDNA clone LD32085; Sprime similar to U31092;
Drosophila melanogaster G protein alpha subunit SmeG-1 (dgqalpha-3)
mRNA, complete cds, mRNA sequence.
ACCESSION AA951538
VERSION AA951538.1 GI:3113775
KEYWORDS EST.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 717)
AUTHORS Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C., Brokstein
P., Lewis,S. and Rubin,G.M.
BDGP/HMMI Drosophila EST Project
Contact: Harvey, D.
Contact: M. Rubin-Molecular and Cell Biology
University of California Berkeley
533 Licko Berkeley, CA 94720-3200, USA
Fax: 510 843 9947
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
Plate: 320, row: H column: 1
High quality sequence stop: 653.

```

```
FEATURES
```

```
Location/Qualifiers
```

```
1..717
```

```
/organism="Drosophila melanogaster"
```

```
/db_xref="taxon:7227"
```

```
/clone_lib="LD32085"
```

```
/sex="male and female"
```

```
/dev_stage="0 to 24 hours mixed stage embryonic"
```

```
/lab_host="X11 Blue"
```

```
/note="Origin: embryo; Vector: pOT2; Site_1: EcoRI; Site_2: XhoI; Sized fractionated cDNAs were directly ligated into pOT2."
```

```
BASE COUNT 217 a 124 c 156 g 320 t
```

```
ORIGIN
```

```
alignment_scores:
```

```
Quality: 723.00      Length: 250
```

```
Ratio: 3.633        Gaps: 2
```

```
Percent Similarity: 79.600      Percent Identity: 55.200
```

```
alignment_block:
```

```
US-08-878-801-2 x AA951538
```

```
Align seg 1/1 to: AA951538 from: 1 to: 717
```

```
124 ValThrThrPheGluLysArgTyrAlaAlaAlaMetGlnTrrLeuTrpAr 140
```

```
4 GTTACCACGTCGAGGATCCATCTTGAATGCCA1CAAAACGCTTTGGGA 53
```

```
140 gAPalaGlyIleArgAlaCysTyrGluArgArgGluGluPheHisLeul 157
```

```
54 CGATGCTGGCATCCAGGATGCTATGATCGTCTAGGGAATATCAGCTGA 103
```

```
157 euAspSerAlaValTyrTyrLeuSerHisLeuGluArgIleThrGluGlu 173
```

```
104 CTGATTCAGCAAAATATATCTGAAGATCTCGATCTGTGTGGCTCAACCT 153
```

```
174 GlyTyrValProThrAlaGlnAspValLeuArgSerArgMetProThrTh 190
```

```
154 GCATATTTACCCACTGACGAAGACATTTTAAGAGTTCGTGTGCCCAAC 203
```

```
190 rGlyIleAsnGluTyrCysPheSerValGlnLysTrAsnLeuArgIleV 207
```

```
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Thu Sep 6 15:16:55 2001

```

204 AGGATAATTGAGTATCCCTTTGATTAGAGAAATCAGATTTAGAATGG 253
207 aAspValGlyGlyGlnLysSerGluArgLysTrpIleHisCysPhe 223
254 TAGACGTCGGTGTGTCAGCGATCCGAGAGAAAGTGGATTTCATTCCTT 303
224 GluAsnValIleAlaLeuIleTyrLeuAlaSerLeuSerGluTyrAspG 240
304 GAGAAATGTGACATCAATATATATTTTGGTAGCGCTATCGGAGTACGATCA 353
240 nCysLeuGluGluAsnGlnGluAsnArgMetLysGlnSerLeuAlaL 257
354 AATCTTGTGTAATCGAATCAATGAGAAATCGAGGAATCTAAAGCTT 403
257 euPheGlyThrIleLeuGluLeuProTyrPheLysSerThrSerValIle 273
404 TATTCGTACTATTAATACATACCTTGGTTTCAAAATTCGTCAGTTATT 453
274 LeuPheLeuAsnLysThrAspIleLeuGluGluLysIleProThrSerHi 290
454 CTTTTCCTGAATAAGAAGGACTTGTGGAAGAGAGAAATAATGTATTGCA 503
290 sLeuAlaThrTyrPheProSerPheGlnGlyProLysGlnAspAlaGluA 307
504 TTTGGTAGACTATTTTCTGAATACGATGGTCTCCGCGAGATGCAATAA 553
307 laAlaLysArgPheIleLeuAspMetTyrThrArgMetTyrThrGlyCys 323
554 CGGCCCGAGAGTTTATCTCGGAATGTTGTAGATTTA..... 591
324 ValAspGlyProGluGlySerLysLysGlyAlaArgSerArgArgLeuPh 340
592 .....AATCCAGATTCGGAANA.....ATTATCTA 617
340 eSerHisTyrThrCysAlaThrAspThrGlnAsnIleArgLysValPheL 357
618 TTCTCATTTTCACGTGTGCTACAGATACGGAATAATATAAGTTTGTGTTG 667
357 ysAspValArgAspSerValLeuAlaArgTyrLeuAspGluIleAsnLeu 373
668 CAGCTGTTAAGGACACAATTCTGCAATCGAACCTTAAGGAATATAAATTG 717

```

